








(TM)

Release 2.1D John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.  
Distribution rights by IntelliGenetics, Inc.

MPearch<sub>pp</sub> protein - protein database search, using Smith-Waterman algorithm

Run on: Mon Apr 1 11:02:14 1996; MasPar time 5.08 Seconds

Tabular output not generated.

Title: >US-08-347-748-2

Perfect Score:

Sequence: 1 MELTELLVMLLTARLT.....PTSP LINTSYTHSQNLSDG 353

### Scoring table:

Gap 11

Searched: 62355 seqs, 7230759 residues

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database: a-geneseq20

1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7  
8:part8 9:part9 10:part10 11:part11 12:part12

Statistics: Mean 34.391; Variance 182.424; scale 0.189

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result	No.	Score	Query Match	Length	DB	ID	Description	Pred. No
	146	5.9	193	2	R11893	Erythropoietin analog	2_41e-02	
	2	144	5.8	193	2	R11891	Erythropoietin analog	3_26e-02
	3	141	5.7	188	3	P60599	Clone lambda HEPORL6	5_15e-02
	4	141	5.7	188	1	P81195	Erythropoietin encode	5_15e-02
	5	141	5.7	193	12	R65499	Human prepro-erythrop	5_15e-02
	6	141	5.7	193	3	P60597	Clone lambda HEPORL3	5_15e-02
	7	141	5.7	193	2	R11892	Erythropoietin analog	5_15e-02
	8	141	5.7	193	3	P50300	Human erythropoietin	5_15e-02
	9	141	5.7	193	2	P70256	Sequence of human ery	5_15e-02
	10	141	5.7	330	4	R23076	Epo:IL-3 short, recom	5_15e-02
	11	141	5.7	349	4	R23079	Epo:IL-3 Flex, recom	5_15e-02
	12	140	5.6	193	3	P60598	Open reading frame co	5_99e-02
	13	140	5.6	193	2	R11858	Modified human erythr	5_99e-02

14	134	5.4	193	3	P50343	EPO encoded by clone	1.47e+01
15	127	5.1	192	3	P50301	Monkey erythropoietin	4.17e-01
16	128	5.1	193	2	R11859	Erythropoietin analog	3.60e-01
17	126	5.1	503	6	R34127	Anexin XI type I fco	4.83e-01
18	121	4.9	166	12	R63647	Mouse erythropoietin.	1.00e+00
19	119	4.8	302	4	R23596	Recombinant hematopoi.	1.34e+00
20	119	4.7	314	4	R23075	IL-3:Epo short, recom	1.34e+00
21	116	4.8	321	4	R23597	Recombinant hematopoi	2.07e+00
22	116	4.7	340	4	R23078	IL-3:Epo Flex, recomb	2.07e+00
23	115	4.6	77	1	P81196	Erythropoietin encode	2.39e+00
24	114	4.6	168	12	R63650	Pig erythropoietin.	2.76e+00
25	115	4.6	3144	11	R58777	Protein encoded by Hu	2.39e+00
26	113	4.5	165	12	R63645	Monkey erythropoietin	3.18e+00
27	113	4.5	166	12	R63644	Human erythropoietin.	3.18e+00
28	113	4.5	166	2	P70398	Sequence of human ery	3.18e+00
29	113	4.5	166	4	R23593	Recombinant hematopoi	3.18e+00
30	113	4.5	167	3	P50299	Human recombinant ery	3.18e+00
31	113	4.5	167	3	P50298	Human recombinant ery	3.67e+00
32	112	4.5	167	2	R07665	Cysteine-added varian	3.67e+00
33	113	4.5	303	4	R23598	Recombinant hematopoi	3.18e+00
34	113	4.5	322	4	R23599	Recombinant hematopoi	3.18e+00
35	111	4.5	334	3	P83194	Sequence of a bioadhe	4.23e+00
36	109	4.4	165	12	R63648	Rat erythropoietin.	5.61e+00
37	109	4.4	505	6	R56664	Anexin XI type I fco	5.61e+00
38	109	4.4	610	9	R56664	Mutant platelet glyco	5.61e+00
39	109	4.4	610	9	R51116	Platelet glycoprotein	5.61e+00
40	106	4.3	165	12	R63646	Monkey erythropoietin	8.55e+00
41	107	4.3	166	12	R62603	Cat erythropoietin.	7.43e+00
42	107	4.3	439	5	R28150	Sugar beet chitinase	7.43e+00
43	106	4.3	1319	8	R47043	Mammalian son of seve	8.55e+00
44	105	4.2	742	3	R14411	Human milk bile-salt	9.83e+00
45	105	4.2	744	8	R45189	BSTL/CTL.	9.83e+00

## ALIGNMENTS

RESULT	1
ID	RI1893 standard; Protein; 193 AA.
AC	RI1893;
DT	22-JUL-1991 (first entry)
DE	Erythropoietin analogue [Pro]124,[Thr]125.
KW	Human erythropoietin; EPO; Isoform; analogue; haematocrit;
KW	sialic acids; glycosylation.
OS	Homo sapiens.
PN	W09105867-A.
PD	EP-428267-A.
PE	02-MAY-1991.
PF	09-OCT-1990; U05758.
PR	13-OCT-1989; US-421444.
PA	(AMGE-) AMGEN INC.
PI	Strickland TW, Byrne TE, Elliott SG;
DR	WPI; 91-148745/20.
PT	Recombinant erythropoietin iso-forms and purificn. - increase
PT	haematocrit levels in mammals and contg. specific number of
PT	sialic acids
PS	Claim 33; Page 45; 60pp; English
CC	The analogue was constructed to add an O-glycosylation site
CC	at Thr125 (amino acid 154 in this sequence) and was produced by
CC	expression of DNA obtained by site-directed mutagenesis of DNA
CC	encoding EPO. The analogue has a higher sialic acid content than
CC	human EPO and has increased biological activity.
CC	It can be used to treat mammals to cause bone marrow cells to
CC	increase prodn. of reticulocytes and red blood



PA (INTE-) Integrated genetics  
PI Beck AK, Withy RM, Zabrecky JR, Massiello NC;  
DR WP1; 88-134531/20.  
DR N-PSDB; N81554.  
PT Recombinant human erythropoietin - produced by a transformed rodent  
PT capable of producing N-linked and O-linked glycosylated human  
PT erythropoietin.  
PS Disclosure; *pf* English.  
CC EPO 104B was one of four positive clones isolated from a cDNA  
-CC library prepd. from mRNA extracted from a human fetus of about 20  
CC wk. gestation. The clone was identified using two probes, EPO1 and  
CC P-8062 based on the published sequence of EPO (Nature (1985) Vol.313,  
CC p.806). The sequence between nucleotides 63 and 724 has 100% hom-  
CC logy with the published sequence. It encodes the 166 AAs of the  
CC mature EPO protein and 22 AAs of the leader sequence. This clone  
CC and a second, EPO 125, were used to construct a full length clone  
CC which was expressed in rodent epithelial cells.  
CC See also P81196.  
SQ Sequence 188 Aa;

[illegible]

	RESULT	5
ID	R65499	standard; Protein; 193 AA.
AC	R65499;	
DT	24-JUN-1995	(first entry)
DE	Human prepro-erythropoietin.	
KW	Erythropoietin; therapeutic; ss.	
OS	Synthetic.	
FH	Key	Location/Qualifiers
FT	Peptide	1..27
FT	/note="leader peptide"	
PN	W09425055-A.	
PD	10-NOV-1994.	
PF	29-APR-1994; U04755.	
PR	29-APR-1993; U5-055076.	
PA	(ABBO ) ABBOTT LAB.	
P1	Devries PJ, Mellowitz BS, Meuth JL, Okasinski GF,	
P1	Schaefer VG.	
DR	WPI; 94-357906/44.	
PT	P-PSDB; Q14760.	
PT	Erythropoietin analogues - useful for treatment of anaemia and have enhanced erythropoietic effect.	
PS	Disclouser; Page 38-39; 56pp; English.	
CC	DNA encoding human prepro-erythropoietin may be ligated into an expression vector for erythropoietin expression in a CHO cell culture. Site-directed mutagenesis may be used in the construction of EPO analogues with improved activity, which may be used in pharmaceutical compositions for inducing erythropoiesis and treating anaemia.	
CC	Sequence 193 AA;	
Q0		

Query Match 5.7%; Score 141; DB 12; Length 193;

Best local similarity 27.6%; Pred. No. 5, 15a-02;  
Matches 27; Conservative 27; Mismatches 41; Indels 3; Gaps 3;

ID RESULT 6  
 AC ID P60597 strand; protein; 193 AA.  
 DT P60597,  
 DT 01-JAN-1980 (first entry)  
 DE Clone lambda HEP0FL3 encoding human erythropoietin.  
 KM Erythropoietin; lambda HEP0FL3; recombinant plasmid vector; anaemia;  
 KM mammalian cell culture; 3T3; CHO; Chinese hamster ovary; ss.  
 KM Homo sapiens.  
 OS M08603520-A.  
 PN 19-JUN-1986.  
 PF 03-DEC-1985; 002405.  
 PR 04-DEC-1984; US-677813.  
 PR 03-JAN-1985; US-688622.  
 PR 22-JAN-1985; US-693258.  
 PA (GENE-) GENETICS INST INC.  
 PA (FRIT/) FRITSCH E.  
 PI Fritsch E, Hewick RM, Jacobs K;  
 DR WPI, 86-169459/26.  
 DR N-PSDB; N60513.  
 PT Prodn. of human cDNA clone expressing erythropoietin - for mass  
 PT prodn. of erythropoietin, useful for treating anaemia  
 PS Disclosure; Page 7; 61pp; English.  
 CC A recombinant plasmid vector expressing this clone is expressed in e.  
 CC q 3T3 or CHO cell cultures. The produced erythropoietin is useful  
 CC for treatment of anaemia, especially renal anaemia. The cloned gene  
 CC expresses high levels of the protein and thus provides a means of  
 CC mass production. See also N60514-21 and P60598-99.  
 QO Sequence 193 AA;

Query Match	5.74;	Score 141;	DB 3;	Length 193;
Best Local Similarity	27.66;	Pred. No. 5.15e-02;		
Matches	27;	Conservative	21;	Mismatches 41; Indels 3; Gaps 3.
Db	12	lllallslplgipvlgagpprlcdsrylevlylleakeenitgtccaeblcslnenitvtpdt	71	
Oy	7	llvvlmlrlarlrt-lsspdpapcdrlvslsllrldshvllahsrllsqcevpnr lppvvlipav	65	
Db	72	kmyfawk-rimevggagvevwtgalllleavrlrgall	108	
Oy	66	dfslgckrtyoml-etkrqddlcaavtlllecvmhaargq	102	

RESULT	7
ID	R11892 standard; Protein; 193 AA.
AC	R11892; 22-00L-1991 (first entry)
DE	Erythropoietin analogue [Thr]125.
KM	Human erythropoietin; EPO; isoform; analogue; haematocrit;
KW	salic acids; glycosylation.
OS	Homo sapiens.
PN	W09105867-A.
PN	EP-428267-A.

Query Match 5.7%; Score 141; DB 12; Length 193;

PD 02-MAY-1991.  
PF 09-OCT-1990; 005758.  
PF 13-OCT-1989; US-421444.  
PA (AMGE-) AMGEN INC.  
PI Strickland TW, Byrne TE, Elliott SG;  
DR WPI; 91-148745/20.  
DR WPI; 91-150265/21.  
PT Recombinant erythropoietin iso-forms and purificn. - increase  
PT haematocrit levels in mammals and contg. specific number of  
-PT static acids  
PS Claim 33; Page 45; 60pp; English  
CC The analogue was constructed to add an O-glycosylation site  
CC at Thr125 (amino acid 154 in this sequences) and was produced by  
CC expression of DNA obtained by site-directed mutagenesis of DNA  
CC encoding Epo. The analogue has a higher static acid content than  
CC human Epo and has increased biological activity.  
CC It can be used to treat mammals to cause bone marrow cells to  
CC increase prodn. of reticulocytes and red blood  
CC thereby increasing haematocrit levels.  
CC See also R11859, R11891-93.  
SQ Sequence 193 Aa;

Query Match	5.7%;	Score 141;	DB 2;	Length 193;
Best Local Similarity	27.6%;	Pred. No. 5.15e-02;		
Matches	27;	Conservative	27;	Mismatches 41; Indels 3; Gaps 3;

```
Dd      12  |||sl||p||g||pv||gaprrl|ldcrv|eryll|eekaeemittgcacahlenitptcd 71
        || : | |   ::| : || | : | : : | : | : | : | : | : | : |
Qy      7  LNVVLLLTARLT-LSSAPPADLDRTVKSLRDSDHYLHRSLSQQPEVNHLPPYVLNAV 65
```

Db	72	kvnlyawk-rtmvgqgawewqglla seavlrqall	108
		::    ::	
		:: ::  ::  ::  ::  ::	
Qy	66	dfslgckrtqme-etrkqddllgavtllllecymaargq	102

RESULT	8
ID	P50300 standard; protein; 193 AA.

DT	01-JAN-1980 (first entry)
DE	Human erythropoietin encoded by positive clone (phage lambda-hel)
DE	isolated from human fetal liver gene bank.

KW disorder; ss; phage lambda-hel; gene bank.  
 OS Homo sapiens.  
 PN M08502610-A.

PF 11-DEC-1984; U02021.  
PR 13-DEC-1983; US-561024.  
PR 21-FEB-1984; US-582185.

PR 30-NOV-1984; US-675298.  
PA (KIRI-) KIRIN-AMGEN INC.  
DR WPI; 85-159229/26.

PT New polypeptide having properties of erythropoietin - is prepd.  
PT by cultivation of transformed eucaryotic or procaryotic host  
PS Disclosure; Page 43; 113pp; English.

CC lambda $\delta$ -Hb1 is essential for red blood cell formation and is used  
CC for the diagnosis and treatment of blood disorders such as anaemia  
CC large amounts of EPO may be obtained using recombinant DNA

CC techniques in contrast to small amounts obtained from plasma  
CC and urine. This sequence is expressed in *E. coli*. See also  
CC N50345-6, N50348-50 and P50298-99, P50301.

**SQ Sequence 193 AA;**

Query Match	5.7%;	Score 141;	DB 3;	length 193;
Best Local Similarity	27.6%;	Pred. No. 5.15e-02;		

Dd 12 Lllslslspjglvlgaprrllicdsrvlyllleakeamittgcacahcnienvtptc 71  
|| : | | | : || : | | : | : : : : : : : : : :  
Qy 7 LAYVMLLTARLE-SSAPRACDRLRVSKLSDSHVHSRLSGQREVNHLPYPLRAV 65

```
Db 72 kmfiawmk-rmevvggavevwglla1lseavlrgqall 108
      :: ||::|| |:::||:|:|::|
QY 66 DFLSLGEMTKQME-ETKQADILCAVTLLEEGVMARQQL 102
```

RESULT	9
ID	P70256 standard; protein; 193 AA.

DT 19-FEB-1991 (first entry)  
DE Sequence of human erythropoietin (EPO).  
KW Renal anaemia therapy; hormone.

	Key	Location/Qualifiers
FH	Peptide	1..27
FT		
FT	/label=SIGNAL	

	Region	81..97
FT	/note="Fragment that probe N70361 is based on" EP-232034-A.	
FT		
PN		

PF 19-JAN-1987; 300399.  
PR 23-JAN-1986; JP-012868.  
PA (SUMO) SUMITOMO CHEM IND KK.

PI Yanagi H, Ogawa I, Okamoto M, Hozumi T, Soga A, Yoshima T,  
PI Tsutsumi M.  
DR WPI, 87-223006/32.

PT Human erythropoietin prodn. - by culturing human cells, esp.  
PT Namalwa cells, transformed with DNA encoding human erythropoietin  
PS Disclosure; Fig 1; 22pp; English.

CC from the erythropoietin-producing human hepatoma cell H<sub>9</sub>-1. The cDNA  
CC library was screened using the probes given in N70361 and N70362. A  
CC plasmid (named as p58-A20) was isolated. The nucleotide sequence of

SQ Sequence 193 AA;  
 Query Match

Query Match 5.7%; Score 141; DB 2; Length 193;  
Best Local Similarity 27.6%; Pred. No. 5.15e-02;  
Matches 27; Conservative 27; Mismatches 41; Indels 3; Gaps 3;

Db 12 lllslslplglpvlgaprrllcdsrvieryllleakeaenltgcgaehcslnentltpdt 71  
 ||: : | | : |: | | | | : : : | : | : : |  
 Qy 7 lllvmltllrtllt-LSSPAPACDLRVSLKLLDSHVLAHSRLSQCFVHNPLETPVLLPAAV 65

```

Db      72 kmfrywkm-rmevvgqavevngqllalseavlryqall 108
      :: ||::| |:::||||:|:::|
Qy      66 DFLSGEMKTYME-ETKQDILGAVTLLLEGVMARQGL 102

```

RESULT	10
ID	R23076 standard; Protein; 330 AA.
AC	R23076;

DT	20-0CT-1992 (first entry)		
DT	Epo:IL-3 short, recombinant haematopoietic molecule.		
KW	Early MDF; late MDF; haematopoiesis; Epo; IL-3; growth factor.		
OS	Homo sapiens.		
FX	Key	Location/Qualifiers	
FT	Peptide	1..27	
FT	/label= sig_peptide		
FT	Protein	28..330	
FT	/label= mat_protein		
PN	M09206116-A.		
PD	16-APR-1992.		
PD	26-SEP-1991; 007053.		
PR	28-SEP-1990; US-589958.		
PA	(ORTH ) ORTHO PHARM CORP.		
PI	Rosen JI;		
DR	WPI; 92-150819/18.		
DR	N-PSDB; Q24282.		
PT	Recombinant haematopoietic molecules useful in treating		
PT	anemia(s) - comprise IL-3 or GM-CSF, EPO, G-CSF, IL-5 or M-CSF		
PT	and has early and later myeloid differentiation activity		
PS	Disclosure; Page 44; 82pp; English.		
CC	The amino acid sequence given is an Epo:IL-3 hybrid growth factor		
CC	derived from a construction formed by ligating the native Epo signal		
CC	sequence and various synthetic oligonucleotides corresponding to Epo		
CC	and IL-3 gene sequences. This hybrid growth factor is a haematopoietic		
CC	molecule which contains at least a portion of an early MDF and at least		
CC	a portion of a late MDF covalently linked. This compound can be used		
CC	to promote haematopoiesis in a patient.		
CC	The bonding of the early and late factors allows a very high conc. of		
CC	late MDF at the surface of a cell which the early MDF is bound. It		
CC	also allows the early MDF to act more specifically to stimulate only		
CC	the desired lineage, thus reducing undesirable effects. These		
CC	compounds are useful for treating anaemias of various origins eg. renal		
CC	failure and AIDS. It is easier to produce and administer one		
CC	recombinant molecule rather than two separate molecules.		
SQ	Sequence 330 AA;		

```

Query Match: 5.7%; Score 141; DB 4; Length 330;
Best Local Similarity 27.6%; Pred. No. 5.15e-02;
Matches 27; Conservative 27; Mismatches 41; Indels 3; Gaps 3;

Db 12 llllellslrlglpvlgspprlcdsrlvlelyllleakeeenltgcacelhslnenitvptd 71
   || : | | | | | | | | | | | | | | : : : | | : | |
Qy 7 lllvmltllrlartl-lsspparpcdlarvlslkldmslvhsrlsgcpevnaplrrpvlvlpav 65

Db 72 kmfyawk-rmevvgqgavevwmqgiallseaavlrgqall 108
   :: || : || | | : : : || : : : ||
Qy 66 dflslgkmtqme-ftrkqodllgavtllllegvmamrqol 102

RESULT 11
ID R23079 standard; Protein; 349 AA.
AC R23079;
DT 20-OCT-1992 (first entry)
DE Epo:IL-3 Flex, recombinant hematopoietic molecule.
KM Early MDF; late MDF; haematopoiesis; Epo; IL-3; Linker;
growth factor.
OS Homo sapiens.
NCBI Key
FT Peptide 1..27 Location/Qualifiers
FT /label= sig_peptide
FT Protein 28..349
FT /label= mat_protein
DN M09206116-A.

```

PD 16-APR-1992.  
PF 26-SEP-1991; U07053.  
PR 28-SEP-1990; US-589958.  
PA (ORTH.) ORTHO PHARM CORP.  
PI Rosen JI;  
DR WP1; 92-150819/18.  
DR N-PSDB; Q24285.  
PT Recombinant haematopoietic molecules useful in treating  
PT anaemia(s) - comprise IL-3 or GM-CSF, EPO, G-CSF, IL-5 or M-CSF  
PS and has early and later myeloid differentiation activity  
PT Disclosure; Page 51; 82pg; English.  
CC The amino acid sequence given is an Epo:IL-3 hybrid growth factor  
CC derived from a construction formed by ligating the native Epo signal  
CC sequence and various synthetic oligonucleotides corresponding to Epo  
CC and IL-3 gene sequences. This molecule is comparable to the sequence  
CC given in R23076 and contains a flexible linker molecule. This hybrid  
CC growth factor is a haematopoietic molecule which contains at least a  
CC portion of an early MDF and at least a portion of a late MDF  
CC covalently linked. This compound can be used to promote haematopoiesis  
CC in a patient.  
CC The bonding of the early and late factors allows a very high conc. of  
CC late MDF at the surface of a cell which the early MDF is bound. It  
CC also allows the early MDF to act more specifically to stimulate only  
CC the desired lineage, thus reducing undesirable effects. These  
CC compounds are useful for treating anaemias of various origins eg. renal  
CC failure and AIDS. It is easier to produce and administer one  
CC recombinant molecule rather than two separate molecules.  
SQ Sequence 349 Aa;

Query Match	5.7%	Score 141;	DB 4;	Length 349;
Best Local Similarity	27.6%	Pred. No. 5.15e-02;		
Matches	27;	Conservative	27;	Mismatches 41; Indels 3; Gaps 3.
Db	12	llllellepdlgprvlgpprlccdcsvryevylleakeasenttgcacncslmenttypdt	71	
Oy	7	lllvmllllarlt-lsspapacdrrvlstklrdshvllshlsdqcevhplptpvllpav	65	
Db	72	kmyfawk-timevgqagvevmgalllseavirgall	108	
Oy	66	dpslgeknktome-etkkaodilcaavtllllegcvmaargal	102	

RESULT	12	
ID	P60598	standard; protein, 193 AA.
AC	P60598;	
DE	01-JAN-1980	(first entry)
DT	Open reading frame coding for the erythropoietin tryptic fragment	
DE	of lambda HEB01.	
KW	Erythropoietin; lambda HEB01; recombinant plasmid vector; anaemia;	
KW	mammal cell culture; 3T3; CHO; Chinese hamster ovary; ss.	
OS	Homo sapiens.	
PN	W08603520-AA.	
PD	19-JUN-1986.	
PF	03-DEC-1985; U02405.	
PR	04-DEC-1984; US-677813.	
PR	03-JAN-1985; US-688622.	
PR	22-JAN-1985; US-693258.	
PA	(GENE-) GENETICS INST INC.	
PA	(FRIT/) FRITSCHE E.	
PI	Fritsch E, Hewick RM, Jacobs K;	
PI	WPI; 86-169459/26.	
DR	N-PSDB; N60518.	
PT	Prodn. of human cDNA clone expressing erythropoietin - for mass	
PT	prodn. of erythropoietin, useful for treating anaemia	

P5 Disclosure: Page 19; 61pp: English.  
 CC Recombinant plasmid vector lambda HEP1 expressing this genomic  
 CC fragment is expressed in e.g. 3T3 or CHO cell cultures. The  
 CC produced erythropoietin is useful for treatment of anaemia.  
 CC especially renal anaemia. The cloned gene expresses high levels  
 CC of the protein and thus provides a means of mass production. See  
 CC also N60513-17, N60519-21 and P60599.  
 SQ Sequence 193 AA;

Query Match	5.6%;	Score 140;	DB 3;	Length 193;
Best Local Similarity	27.6%;	Pred. No. 5,99e-02;		
Matches	27;	Conservative	27;	Mismatches 41; Indels 3; Gaps 3;

[illegible]

RESULT	13	
ID	R11858 standard; protein; 193 AA.	
AC	R11858;	
DT	24-JUL-1991 (first entry)	
DE	Modified human erythropoietin.	
KW	EPO; anaemia.	
OS	Homo sapiens.	
FH	Key	Location/Qualifiers
FT	Protein	28..193
FT	/label= Mature modified EPO	
FT	Misc.difference 51	
FT	/label= May be absent or any AA except Asn	
FT	Misc.difference 65	
FT	/label= May be absent or any AA except Asn	
FT	Misc.difference 110	
FT	/label= May be absent or any AA except Asn	
FN	EP-427189-A.	
PD	15-MAY-1991.	
PF	06-NOV-1990; 121187.	
PR	07-NOV-1989; JP-289143.	
PA	(SNOW ) SNOW BRAND MILK PRODUCTS.	
PI	Yamauchi K, Ueda M;	
DR	WPI; 91-141727/20.	
PT	New human erythropoietin(s) modified at positions 24, 38 and 83 -	
PT	having improved biological activity for stimulating erythrocyte	
PT	differentiation and treating anemia.	
PS	Disclosure; Page 8-9; 14pp; English.	
CC	Modified EPO has a reduced number of stalleic acid binding sites, and	
CC	the galactose residue is not exposed. By stripping away these sugars	
CC	from the glycoprotein, the affinity of the EPO to it's receptor is	
CC	greatly enhanced. Thus, the modified EPO shows a higher activity and	
CC	is useful in the treatment of anemia.	
Sequence	193 AA;	

Query Match	5.6%;	Score 140;	DB 2;	Length 193;
Best Local Similarity	27.6%;	Pred. No. 5.99e-02;		
Matches 27; Conservative		26; Mismatches 42; Indels 3; Gaps 3;		

Db  
12 LLSLSPGLPVLGAPRLLICSDRVLERYLLAEKAEQITTCGAENCSLNEDITYPDT 71  
|| : || | : || : || ||| : : : : || : ||  
7 LLVAMLLTARLT-LSSPAAPACDLRVTSKLINDSHVLSRSQCPEVHPLPDPVLLPAV 65

```
Db 72 kmfyawk-rmevqgqavewqglallseavlrqall 108
    :: || ::| | :: ::|| |::| ::|
QY 66 DFLSGEMKTQME-ETKAQDILIGAVTLLLEGVMARQQL 102
```

RESULT 14  
ID: P50343 standard; Protein; 193 AA.

DT 10-MAR-1992 (first entry)  
DE EPO encoded by clone lambda-HEPOFL13.  
KW Erythropoietin.  
OS Homo sapiens.

PF	PD	PN	FT	FT	FT	FT	FT	FT	FT
27-DEC-1984;	16-AUG-1985.	DK8406107-A.	/label= mature_EPO	protein	/label= signal_sequence_27..193	1..27	Peptide	Key	Location

FR 27-DEC-1983; US-566057.  
FR 04-DEC-1984; US-677813.  
PR 21-DEC-1984; NO-005186.  
PR 21-DEC-1984; ZA-010034.  
PR 22-JAN-1985; US-693258.  
PR 22-JAN-1985; US-688622.  
PA (GENE-) GENETICS INST.  
PI Kaufman RJ.  
DR WPI; 85-318061/51.  
DR N-PSDB; NS0443.  
PT Vector system for introducing heterologous DNA into eukaryotic cells - comprises prod. gene and accessory DNA for enhanced expression of heterologous protein by the cells.  
PT disclosure; Fig 13; 62pp; Danish.  
PS See also US4740461 88.04.26 (8819) (first major country equivalent).  
CC The sequence is encoded by clone lambda-HEPFL3 which contains the gene encoding erythropoietin obt'd. from a human foetal DNA library. The gene can be expressed using a novel vector system disclosed in the specification.  
CC See also P50342.  
CC Sequence 193 Aa;  
Q2

Query Match	5.4%;	Score 134;	DB 3;	length 193;
Best Local Similarity	26.3%;	Pred. No. 1.47e-01;		
Matches	35;	Conservative	38;	Mismatches 54;
				Indels 6;
				Gaps 5;

Db 12 llllellslpqlpvrpaprllcdsryleryllleakgaenitfgcaebclnienllypdt 71  
 Qy 7 llvmlmlfmlarlltssppp--ACDLRVtSKLLRDSHVHfSRtSQCPVEVHPllpVLLPAV 65  
 Db 72 kmfyamk-lveagqaavewqgalllseavlyrgqallvnssqgmepql--hvdkaavg 128  
 Qy 66 DFSLCEKMTQREETR--ADILLGAVTLLLEGVMAARCOLPTCTISLLGQLSGOVRLILGA 1244  
 Db 129 lrsf-ftllrpqg 140  
 Qy 125 lqslmqtqlpPg 137

RESULT	15
ID	P50301 standard; protein; 192 AA.
AC	P50301;
DT	01-JAN-1980 (first entry)



Release 2.1D John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.  
Distribution rights by IntelliGenetics, Inc.

MParch<sub>pp</sub> protein - protein database search, using Smith-Waterman algorithm

Run on: Mon Apr 1 11:01:07 1996; MasPar time 11.78 Seconds

Tabular output not generated.

Title: >US-08-347-748-2

Perfect Score: 2493

Sequence: 1 MELTELLLVMLLTARLT.....PTSP LINTSYTHSONISQEG 353

Scoring table: PAM 150

Gap 11

Searched: 82306 seqs, 25270970 residues

Post-processing: Minimum Match 0%

Database:  pir46

```
1:ann1 2:ann2 3:ann3 4:unann1 5:unann2 6:unann3 7:unann4
8:unann5 9:unann6 10:unann7 11:unann8 12:unrev1
13:unrev2
```

Statistics: Mean 46.752; Variance 133.030; scale 0.351

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description	Pred. No
1	2493	100.0	353	13	S48740	thrombopoietin - hum	0.00e+00
2	2493	100.0	353	13	S48731	c-MET ligand - human	0.00e+00
3	1748	70.1	356	13	S45330	thrombopoietin - mou	3.38e-23
4	1141	68.6	326	13	JC4125	thrombopoietin - rat	1.06e-22
5	1143	45.8	286	10	A55330	megakaryocyte growth	1.06e-14
6	155	6.2	192	5	A24902	erythropoietin precu	3.61e-05
7	151	6.2	192	13	S28148	erythropoietin - rat	4.65e-05
8	144	5.7	193	2	Z0HU	erythropoietin precu	1.19e-03
9	139	5.6	503	2	L0RB11	amexin XI - rabbit	1.94e-03
10	132	5.3	485	8	A33647	erythral surface gly	1.05e-02
11	127	5.1	192	5	J00173	erythropoietin precu	3.41e-02
12	127	5.1	193	5	A24901	erythropoietin precu	3.41e-02

13	126	5.1	503	2	LJB011	anexin XI form A -	4.30e+02
14	121	4.9	350	9	SZ2456	hydroxyproline-rich	1.36e-01
15	121	4.9	505	10	A53152	anexin XI - human	1.36e-01
16	116	4.7	228	12	S53504	extensin-like protei	4.21e-01
17	116	4.7	817	9	S51342	verprolin - yeast (S	4.21e-01
18	115	4.6	297	12	S23737	proline-rich protein	5.26e-01
19	115	4.6	421	13	S29599	Acrosin (EC 3.4.21.1	5.26e-01
20	115	4.6	441	12	S44524	hypothetical protein	5.26e-01
21	115	4.6	441	9	S37881	hypothetical protein	5.26e-01
22	115	4.6	1874	1	J00533	RNA-directed RNA pol	5.26e-01
23	115	4.6	3144	10	A00686	Huntington's disease	5.26e-01
24	112	4.5	267	12	S08314	cell wall glycoprote	1.02e+00
25	112	4.5	328	9	J00985	hydroxyproline-rich	1.02e+00
26	112	4.5	1232	10	S40766	hypothetical protein	1.02e+00
27	109	4.4	268	6	S06666	protein TPX-VI3 - Th	1.95e+00
28	110	4.4	473	8	S50755	hypothetical protein	1.57e+00
29	109	4.4	474	6	S15921	protein TPX-VI3 - Th	1.95e+00
30	109	4.4	505	2	S23447	anexin XI form B -	1.95e+00
31	109	4.4	626	3	NBHUIA	platelet glycoprotei	1.95e+00
32	106	4.3	153	9	S48998	hypothetical protein	3.71e+00
33	107	4.3	283	9	S13363	hydroxyproline-rich	3.00e+00
34	106	4.3	353	11	S37186	transcription factor	3.71e+00
35	106	4.3	377	11	C56288	hepatocyte nuclear f	3.71e+00
36	106	4.3	373	10	A48018	muchin, M2=low molec	3.71e+00
37	107	4.3	410	7	B85523	ferredoxin reductase	3.00e+00
38	106	4.3	434	12	S11967	module-specific (hyd	3.71e+00
39	107	4.3	439	8	S45055	chitinase (EC 3.2.1.	3.00e+00
40	107	4.3	464	11	A47655	apiliosome-associated	3.00e+00
41	106	4.3	593	13	S44555	glycoprotein C (thom	3.71e+00
42	106	4.3	1206	13	S24407	formin isoform IV -	3.71e+00
43	106	4.3	1336	11	S21391	hypothetical protein	3.71e+00
44	106	4.3	1336	11	S23716	hypothetical protein	3.71e+00
45	106	4.3	1468	13	S11515	formin - mouse	3.71e+00

## ALIGNMENTS

RESULT	1
ENTRY	S58740
TITLE	thrombopoietin - human
ORGANISM	#formal name Homo sapiens #common name man
DATE	01-Aug-1995 #sequence_revision 01-Aug-1995 #text_change 01-Aug-1995

## REFERENCE

**Authors**  
Sohma, Y.; Akahori, H.; Seki, N.; Hori, T.; Ogami, K.; Kato  
T.; Shimada, Y.; Kawamura, K.; Miyazaki, H.

## # journa

**#title** Molecular cloning and chromosomal localization of the human thrombopoietin gene.

**Access**

```
##status preliminary
##residues 1-353 ##label 50H
```

## SUMMARY

Query Match	100.0%;	Score 2493;	DB 13;	Length 353;
Best Local Similarity	100.0%;	Prod No 0	000+00;	

DEBC LOC  
Matches

```
1 me[te]]vm|tar|sannac|rv|sk|rdsbv|sr|sacdev|p|ov 60
          00, collectvalve 0, mismatches 0, inlets 0, caps 0
```

5

1 MELTELLVMIITARTLTSSPAPACDLRLSKLDRSDHLSRLSQCEVHP LPTPV 600

## Db 61

1. *Chlorophyll a* (Chl *a*)  
 2. *Chlorophyll b* (Chl *b*)  
 3. *Chlorophyll c* (Chl *c*)  
 4. *Chlorophyll d* (Chl *d*)  
 5. *Chlorophyll e* (Chl *e*)  
 6. *Chlorophyll f* (Chl *f*)  
 7. *Chlorophyll g* (Chl *g*)  
 8. *Chlorophyll h* (Chl *h*)  
 9. *Chlorophyll i* (Chl *i*)  
 10. *Chlorophyll j* (Chl *j*)  
 11. *Chlorophyll k* (Chl *k*)  
 12. *Chlorophyll l* (Chl *l*)  
 13. *Chlorophyll m* (Chl *m*)  
 14. *Chlorophyll n* (Chl *n*)  
 15. *Chlorophyll o* (Chl *o*)  
 16. *Chlorophyll p* (Chl *p*)  
 17. *Chlorophyll q* (Chl *q*)  
 18. *Chlorophyll r* (Chl *r*)  
 19. *Chlorophyll s* (Chl *s*)  
 20. *Chlorophyll t* (Chl *t*)  
 21. *Chlorophyll u* (Chl *u*)  
 22. *Chlorophyll v* (Chl *v*)  
 23. *Chlorophyll w* (Chl *w*)  
 24. *Chlorophyll x* (Chl *x*)  
 25. *Chlorophyll y* (Chl *y*)  
 26. *Chlorophyll z* (Chl *z*)  
 27. *Chlorophyll aa* (Chl *aa*)  
 28. *Chlorophyll ab* (Chl *ab*)  
 29. *Chlorophyll ac* (Chl *ac*)  
 30. *Chlorophyll ad* (Chl *ad*)  
 31. *Chlorophyll ae* (Chl *ae*)  
 32. *Chlorophyll af* (Chl *af*)  
 33. *Chlorophyll ag* (Chl *ag*)  
 34. *Chlorophyll ah* (Chl *ah*)  
 35. *Chlorophyll ai* (Chl *ai*)  
 36. *Chlorophyll aj* (Chl *aj*)  
 37. *Chlorophyll ak* (Chl *ak*)  
 38. *Chlorophyll al* (Chl *al*)  
 39. *Chlorophyll am* (Chl *am*)  
 40. *Chlorophyll an* (Chl *an*)  
 41. *Chlorophyll ao* (Chl *ao*)  
 42. *Chlorophyll ap* (Chl *ap*)  
 43. *Chlorophyll aq* (Chl *aq*)  
 44. *Chlorophyll ar* (Chl *ar*)  
 45. *Chlorophyll as* (Chl *as*)  
 46. *Chlorophyll at* (Chl *at*)  
 47. *Chlorophyll au* (Chl *au*)  
 48. *Chlorophyll av* (Chl *av*)  
 49. *Chlorophyll aw* (Chl *aw*)  
 50. *Chlorophyll ax* (Chl *ax*)  
 51. *Chlorophyll ay* (Chl *ay*)  
 52. *Chlorophyll az* (Chl *az*)  
 53. *Chlorophyll aza* (Chl *aza*)  
 54. *Chlorophyll abz* (Chl *abz*)  
 55. *Chlorophyll acz* (Chl *acz*)  
 56. *Chlorophyll adz* (Chl *adz*)  
 57. *Chlorophyll aez* (Chl *aez*)  
 58. *Chlorophyll afz* (Chl *afz*)  
 59. *Chlorophyll agz* (Chl *agz*)  
 60. *Chlorophyll ahz* (Chl *ahz*)  
 61. *Chlorophyll aiz* (Chl *aiz*)  
 62. *Chlorophyll ajz* (Chl *ajz*)  
 63. *Chlorophyll akz* (Chl *akz*)  
 64. *Chlorophyll alz* (Chl *alz*)  
 65. *Chlorophyll amz* (Chl *amz*)  
 66. *Chlorophyll anz* (Chl *anz*)  
 67. *Chlorophyll aoz* (Chl *aoz*)  
 68. *Chlorophyll apz* (Chl *apz*)  
 69. *Chlorophyll aqz* (Chl *aqz*)  
 70. *Chlorophyll arz* (Chl *arz*)  
 71. *Chlorophyll asz* (Chl *asz*)  
 72. *Chlorophyll atz* (Chl *atz*)  
 73. *Chlorophyll auz* (Chl *auz*)  
 74. *Chlorophyll avz* (Chl *avz*)  
 75. *Chlorophyll awz* (Chl *awz*)  
 76. *Chlorophyll axz* (Chl *axz*)  
 77. *Chlorophyll ayz* (Chl *ayz*)  
 78. *Chlorophyll ayz* (Chl *ayz*)  
 79. *Chlorophyll azz* (Chl *azz*)  
 80. *Chlorophyll azaa* (Chl *aza*)  
 81. *Chlorophyll abz* (Chl *abz*)  
 82. *Chlorophyll acz* (Chl *acz*)  
 83. *Chlorophyll adz* (Chl *adz*)  
 84. *Chlorophyll aez* (Chl *aez*)  
 85. *Chlorophyll afz* (Chl *afz*)  
 86. *Chlorophyll agz* (Chl *agz*)  
 87. *Chlorophyll ahz* (Chl *ahz*)  
 88. *Chlorophyll aiz* (Chl *aiz*)  
 89. *Chlorophyll ajz* (Chl *ajz*)  
 90. *Chlorophyll akz* (Chl *akz*)  
 91. *Chlorophyll alz* (Chl *alz*)  
 92. *Chlorophyll amz* (Chl *amz*)  
 93. *Chlorophyll anz* (Chl *anz*)  
 94. *Chlorophyll aoz* (Chl *aoz*)  
 95. *Chlorophyll apz* (Chl *apz*)  
 96. *Chlorophyll aqz* (Chl *aqz*)  
 97. *Chlorophyll arz* (Chl *arz*)  
 98. *Chlorophyll asz* (Chl *asz*)  
 99. *Chlorophyll atz* (Chl *atz*)  
 100. *Chlorophyll auz* (Chl *auz*)  
 101. *Chlorophyll avz* (Chl *avz*)  
 102. *Chlorophyll awz* (Chl *awz*)  
 103. *Chlorophyll axz* (Chl *axz*)  
 104. *Chlorophyll ayz* (Chl *ayz*)  
 105. *Chlorophyll ayz* (Chl *ayz*)  
 106. *Chlorophyll azz* (Chl *azz*)  
 107. *Chlorophyll azaa* (Chl *aza*)  
 108. *Chlorophyll abz* (Chl *abz*)  
 109. *Chlorophyll acz* (Chl *acz*)  
 110. *Chlorophyll adz* (Chl *adz*)  
 111. *Chlorophyll aez* (Chl *aez*)  
 112. *Chlorophyll afz* (Chl *afz*)  
 113. *Chlorophyll agz* (Chl *agz*)  
 114. *Chlorophyll ahz* (Chl *ahz*)  
 115. *Chlorophyll aiz* (Chl *aiz*)  
 116. *Chlorophyll ajz* (Chl *ajz*)  
 117. *Chlorophyll akz* (Chl *akz*)  
 118. *Chlorophyll alz* (Chl *alz*)  
 119. *Chlorophyll amz* (Chl *amz*)  
 120. *Chlorophyll anz* (Chl *anz*)  
 121. *Chlorophyll aoz* (Chl *aoz*)  
 122. *Chlorophyll apz* (Chl *apz*)  
 123. *Chlorophyll aqz* (Chl *aqz*)  
 124. *Chlorophyll arz* (Chl *arz*)  
 125. *Chlorophyll asz* (Chl *asz*)  
 126. *Chlorophyll atz* (Chl *atz*)  
 127. *Chlorophyll auz* (Chl *auz*)  
 128. *Chlorophyll avz* (Chl *avz*)  
 129. *Chlorophyll awz* (Chl *awz*)  
 130. *Chlorophyll axz* (Chl *axz*)  
 131. *Chlorophyll ayz* (Chl *ayz*)  
 132. *Chlorophyll ayz* (Chl *ayz*)  
 133.



```
|||||
Qy 61 LHPADNSLGEKKTQMEETKADIIIGANTLILEGVMARQGLPTCTSLSLQSLSQVRL 120
Db 121 lllqalqslqtqqlppqgrrtahnkdnaifllefqhlllrqkvrflmlvggsclvrrappt 180
|||
Qy 121 LIGALQSLTCTQLPPOGRTTAHKDNPNAIFLSFOHLRGRVRFMLVGSSTLCVRARAPT 180
Db 181 avpsetalvltlnelprntegllletnfaartgsgllkwgggfrakipqllnctfrel 240
|||
Qy 181 AVPSRSTLVLTINELPNRTSGILETNFTASARTSGSLMKQOGFRAKIPGLNQTSSSL 240
Db 241 dqipylnrlthelngtrqqlfpgprrtlqapdisqetdsgslpnnlqpygsppptpp 300
Qy 241 DQIPCYLNRHIELNCTRGLFPGPSRRTLCAPDISSGTSYDGLSPNLDQPSYSPPTHP 300
Db 301 tqgytlfplppltpcpvqqlhpllpdpsaptpcptpplntaythsgnlsqeg 353
|||
Qy 301 TQGYTLFPLPPTLPPTPVQQLHP LHPDPSAPTPPTSPPLMNTSYTHSQNLSDQEG 353

RESULT 2
ENTRY 2
TITLE 545331 #type complete
ORGANISM c-MPL ligand - human
#formal_name Homo sapiens #common_name man
DATE 10-Dec-1994 #sequence_revision 10-Dec-1994 #text_change
10-Dec-1994

ACCESSIONS
545331
REFERENCE
#authors
de Sauvage, F.J.; Haes, P.E.; Spencer, S.D.; Malloy, B.E.;
Gurney, A.L.; Spencer, S.A.; Darbonne, W.C.; Henzel, W.J.;
Wong, S.C.; Kuang, M.J.; Olee, K.J.; Hultgren, B.; Solberg
Jr., L.A.; Goeddel, D.V.; Eaton, D.L.
#journal
Nature (1994) 369:533-538
#title
Stimulation of megakaryocytopoiesis and thrombopoiesis by the
c-Mpl ligand.
#accession
545331
#status
preliminary
#residues 1-353 ##label DEA
SUMMARY #length 353 #molecular-weight 37822 #checksum 1597

Query Match 100.0%; Score 2493; DB 13; Length 353;
Best Local Similarity 100.0%; Pred. No. 0.00e+00;
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Db 301 tqgytlfplppltpcpvqqlhpllpdpsaptpcptpplntaythsgnlsqeg 353
|||||
Qy 301 TQGYTLFPLPPTLPPTPVQQLHP LHPDPSAPTPPTSPPLMNTSYTHSQNLSDQEG 353

RESULT 3
ENTRY 3
TITLE 545330 #type complete
ORGANISM thrombopoietin - mouse
#formal_name Mus musculus #common_name house mouse
DATE 10-Dec-1994 #sequence_revision 10-Dec-1994 #text_change
10-Dec-1994

ACCESSIONS
545330
REFERENCE
#authors
Lok, S.; Kaushansky, K.; Holly, R.D.; Kuiper, J.L.;
Lofton-Day, C.E.; Oort, P.J.; Grant, F.J.; Heipel, M.D.;
Burkhead, S.K.; Kramer, J.W.; Bell, L.A.; Sprecher, C.A.;
Blumberg, H.; Johnson, R.; Prunkard, D.; Ching, A.F.T.;
Mathewes, S.L.; Bailey, M.C.; Forstrom, J.W.; Buddle, M.M.;
Osborn, S.G.; Evans, S.J.; Sheppard, P.O.; Pressnell, S.R.;
O'Hara, P.J.; Hagen, F.S.; Roth, G.J.; Foster, D.C.
#journal
Nature (1994) 369:565-568
#title
Cloning and expression of murine thrombopoietin cDNA and
stimulation of platelet production in vivo.
#accession
545330
#status
preliminary
#residues 1-356 ##label LOK
SUMMARY #length 356 #molecular-weight 37835 #checksum 9983

Query Match 70.1%; Score 1748; DB 13; Length 356;
Best Local Similarity 71.6%; Pred. No. 3.38e-23;
Matches 255; Conservative 47; Mismatches 49; Indels 5; Gaps 4;
```

```
Db 1 meltdlllaamlavarltlseppapacdrlvlsklldshvlnherlseqevhpltpfv 60
|||||
Qy 1 MELTELLLVVMLLTARLTLSPPAPACDRLVLSKLSDSHVLSRSLSCQCEVHP LPTPV 60
Db 181 avpsetalvltlnelprntegllletnfaartgsgllkwgggfrakipqllnctfrel 240
|||||
Qy 181 AVPSRSTLVLTINELPNRTSGILETNFTASARTSGSLMKQOGFRAKIPGLNQTSSSL 239
Db 241 pyqslgylnrthpvrngthqlfaqtcltleasdispgafnqslafnldqglpssala 300
|||||
Qy 240 LDQIPCYLNRHIELNCTRGLFPGPSRRTLCAPDISSGTSYDGLSPNLDQPSYSPPTHP 299

Db 301 pdqqltpfppspalptcthsqppqqlhplfpdpsaptpcptpplntaythsgnlsqeg 355
|||
Qy 300 PTQGYTLFPLPPTLPPTPVQQLHP LHPDPSAPTPPTSPPLMNTSYTHSQNLSDQEG 352
```

```
RESULT 4
ENTRY 4
TITLE Jc4125 #type complete
ORGANISM thrombopoietin - rat
#formal_name Rattus norvegicus #common_name Norway rat
DATE 27-Aug-1995 #sequence_revision 27-Aug-1995 #text_change
27-Aug-1995
ACCESSIONS
Jc4125
```

REFERENCE	JC4125
#authors	Ogami, K.; Shinada, Y.; Sohma, Y.; Akahori, H.; Kato, T.; Kawamura, K.; Miyazaki, H.
#journal	Gene (1995) 158:309-310
#title	The sequence of a rat cDNA encoding thrombopoietin.
#accession	JC4125
##status	preliminary
##residues	1-326 ##label OGA
SUMMARY	##cross-references DDBJ:D32207 #length 326 #molecular-weight 34556 #checksum 730
Query Match	68.6%; Score 1711; DB 13; Length 326; Best Local Similarity 76.0%; Pred. No. 1,08e-226;
Matches	241; Conservative 43; Mismatches 31; Indels 2; Gaps 2,
Db	1 meltedllvaallltacltsspppacdrlrlnklldsyllhsrlscqcdmplejpy 60      :     :     :     :     :     :     :     :     :     :   QY 1 MELTELLLVMLLTARLTFLSSPPAPACDRLVSLTSLDSWTLHSLSCQEVNPLPTV 60
Db	61 llpaavfslgwetqtgegakadlllgavalllegymaargqlpepcslaligqlsgvrl 120      :     :     :     :     :     :     :     :     :     : QY 61 LLPAVFSLGEMQMKEETKADLIGAVTLLEGYMARQLPRTLSLLGLSQVAL 120
Db	121 llgaqlglgtcqpqrpggrttakdpsealfelsgqlirkrvflllwepalcrrtlpt 180      :     :     :     :     :     :     :     :     :     : QY 121 LLGAQLGLGTQLP PGRRTAKDPNAIFLTFQHLLRKRVAFLTWGSCTLCVRRAPT 180
Db	181 avpsrtsglllnkfmrtsqllletnfvartagpqlmlnrlgfaklipyqlngtqs 240      :     :     :     :     :     :     :     :     :     : QY 181 AVPSRTSVLTFLNELPRRTSGILENFRTASARTGSLGMKGQGPAAKL -PGLNQTGNS 239
Db	241 lddipynlgnthepvnqhtqlfaqtstqltleapdvvpafnkgsrlphlqglpjpisla 300      :     :     :     :     :     :     :     :     :     : QY 240 LDDIPGYNLRIHELINMGTRTGLEPPERSRRTLGNARDISSTSDTSGLPNLDGYSFSPTNP 299
Db	301 adg-ytlfrpsepftrp 316 :   :     :     :     :     :     :     :     :     :     : QY 300 PTQGTYLFRPEPTLPTP 316
RESULT	5 A55530 #type complete
ENTRY	megakaryocyte growth and development factor, long form - human
TITLE	MPL ligand, long form
ALTERNATE_NAMES	#formal name Homo sapiens #common name man
ORGANISM	20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 31-Mar-1995
DATE	A55530
ACCESSIONS	A55530
REFERENCE	A55530
#authors	Chang, M.-J; McNinch, J., Basu, R.; Shutter, J.; Hsu, R.; Perkins, C.; Mar, V.; Suggs, S.; Welcher, A.; Li, L.; Lu, H.; Bartley, T.; Hunt, P.; Martin, F.; Samal, B.; Bogenberger, J. J. Biol. Chem. (1995) 270:511-514
#journal	Cloning and characterization of the human megakaryocyte growth and development factor (MDF) gene.
#title	A55530
#accession	preliminary; not compared with conceptual translation
##status	DNA
##molecule_type	1-286 ##label CHA
##residues	
##cross-references	CB:U01701
GENETICS	MGDF
. #gene	

```

#map_position 3q26.3
KEYWORDS      alternative splicing; cytokine
SUMMARY       #length 286 #molecular-weight 31544 #checksum 6126

Query Match      45.8%; Score 1143; DB 10; length 286;
Best Local Similarity 94.2%; Pred. No. 1,06e-142;
Matches 163; Conservative 2; Mismatches 6; Indels 2; Gaps 2;

Db 1 melteillyvmltlarllssppacdlrvlakllrghvlnharlegscvnhp1ptpv 60
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 1 MELTEILLYVWMLLTARLLTSSPAPACDRLVLSKLRSHVLSRLSQCPENVH1PTPV 60
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

Db 61 llpavfslgskwtqmeekaqdl1gavtlllegvmaargqlptc1s6llsglsgvrr1 120
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 61 LLPAVDFS1GSKMTQMEETKAKDILGAVTLLLEGVMAARGQLPTC1SL1SGVRR1 120
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

Db 121 llgalgsllgfc1ppqgrtrtkhknaf1lscfpl1lrqkd-fwi-yogdk1hnc1 171
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 121 LLGALGSLLGFTQPPGRTTRTKHKNAP1LSCFPL1LRQKVF1WLVGSGTLCV 173
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 6
ENTRY      A24902 #type complete
TITLE      erythropoietin precursor - mouse
ORGANISM   Mus musculus #common_name house mouse
DATE       25-Oct-1987 #sequence_revision 25-Oct-1987 #text_change 18-Jun-1993
ACCESSIONS A24902
REFERENCE   A24902
#authors    Shoemaker, C.B.; Mitscock, L.D.
#journal     Mol. Cell. Biol. (1986) 6:849-858
#title       Murine erythropoietin gene: cloning, expression, and human
              gene homology.
#cross-references M01D:87039105
#accession   A24902
#molecule_type DNA
#residues    1-192 ##label SHO
#note         the authors translated the codon TTA for residue 12 as
              Phe, TTA for residue 43 as Phe and ATG for residue 145
              as Ile

GENETICS
#introns     5/1; 52/3; 81/3; 141/3
CLASSIFICATION #superfamily erythropoietin
KEYWORDS      erythropoiesis; glycoprotein; hormone
SUMMARY       #length 192 #molecular-weight 21339 #checksum 6155

Query Match      6.2%; Score 155; DB 5; length 192;
Best Local Similarity 26.8%; Pred. No. 3.61e-05;
Matches 34; Conservative 35; Mismatches 54; Indels 4; Gaps 4;

Db 10 llllsl1llp1glpvlcapp1lclclsr1v1leakeaanvtmgcaegpr1semit1vad 69
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 6 LLLVWMLLTARLLT-LSSPAPACDRLVLSKLRSHVLSRLSQCPENVH1PTPV1LPA 64
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

Db 70 tkynfyawk-rmevegea1evwqgl1seall1gaqall1ansegpc1qlh1dk1a1eq1 128
:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 65 VDS1GCEKMTQME-ETKAKDILGAVTLLLEGVMAARGQLPTC1SL1SGVRR1-L 122
:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

Db 129 reltail 135
::|1111
Qy 123 GAGQSL 129

```

```

ENTRY      S28148      #type complete
TITLE      erythropoietin - rat
ORGANISM   #formal name Rattus norvegicus #common name Norway rat
DATE       22-Nov-1993 #sequence_revision 22-Nov-1993 #text_change
22-Nov-1993

ACCESSIONS
S28148
REFERENCE
#authors    Nagao, M.; Suga, H.; Okano, M.; Masuda, S.; Narita, H.;
            Ikura, K.; Sasaki, R.
#journal    Biochim. Biophys. Acta (1992) 1171:99-102
#title      Nucleotide sequence of rat erythropoietin.
#accession  S28148
#status     preliminary
#residues   1-192 ##label NAG
SUMMARY     #length 192 #molecular-weight 21286 #checksum 4587

Query Match      6.2%; Score 154; DB 13; Length 192;
Best Local Similarity 25.8%; Pred. No. 4.65e-05;
Matches 34; Conservative 35; Mismatches 58; Indels 5; Gaps 5;

Db 10 LILLELLEPLGLPVLGAPRLICDSRYLYLEAKAENITGCAHCELNITPTD 69
    ||| : ||| : | : | : || ||| : : : : | : | : | :
Oy 6 LLVYMLLTARLT-LSSPAPACDLRYLSKLIRDSHYLSRISQCPVHPLETPVLLPA 64

Db 70 tkyvfwkmkveeqavevwqglalleaillqagallqansgppeslqlhikaisqlr 129
    :: || : | : | : : ||| || : : : | : | : | :
Oy 65 VDFSLGEMKTQMEYKADILGAVTLLLEGVMAARG-QLGRT-CLSSLLQGLSGQVR-LL 121

Db 130 sltstllrvlqag 141
    | : | : ||| : |
Oy 122 -LGAAGSLGRTQ 132

RESULT      8
ENTRY      20HU      #type complete
TITLE      erythropoietin precursor - human
ORGANISM   #formal name Homo sapiens #common name man
DATE       27-Nov-1985 #sequence_revision 27-Nov-1985 #text_change
22-Apr-1995

ACCESSIONS
A01855; A24744; A25384; A22210
REFERENCE
A01855
#authors    Jacobs, K.; Shoemaker, C.; Rudersdorf, R.; Neill, S.D.;
            Kaufman, R.J.; Mufson, A.; Sehara, J.; Jones, S.S.; Hewick,
            R.; Fritsch, E.F.; Kawakita, M.; Shimizu, T.; Miyake, T.
            Nature (1985) 313:806-810
#journal    Isolation and characterization of genomic and cDNA clones of
            human erythropoietin.
#title      human erythropoietin.
#cross-references M01D:85137899
#accession  A01855
#molecule_type mRNA; DNA
#residues   1-193 ##label JAC
##cross-references GB:X02157; GB:X02158
REFERENCE
A24744
#authors    Lin, F.K.; Suggs, S.; Lin, C.H.; Browne, J.K.; Smalting, R.;
            Egrie, J.C.; Chen, K.K.; Fox, G.M.; Martin, F.; Stabinsky,
            Z.; Badrawi, S.M.; Lai, P.H.; Goldwasser, E.
            Proc. Natl. Acad. Sci. U.S.A. (1985) 82:7580-7584
#journal    Cloning and expression of the human erythropoietin gene.
#title      Cloning and expression of the human erythropoietin gene.
#cross-references M01D:86067948
#accession  A24744
#molecule_type DNA
#residues   1-193 ##label LIN
##cross-references GB:M11319
REFERENCE
A25384

```

```

#authors    Lai, P.H.; Everett, R.; Wang, F.F.; Arakawa, T.; Goldwasser,
            E.
#journal    J. Biol. Chem. (1986) 261:3116-3121
#title      Structural characterization of human erythropoietin.
#cross-references M01D:86140080
#accession  A25384
#molecule_type protein
#residues   28-86,'Q',87-193 ##label LAI
#experimental source urine
#note        forms without the carboxyl-terminal residue and the four
            carboxyl-terminal residues were observed

REFERENCE
A22210
#authors    Yanagawa, S.; Hiraide, K.; Ohnota, H.; Sasaki, R.; Chiba, H.;
            Ueda, M.; Goto, M.
#journal    J. Biol. Chem. (1984) 259:2707-2710
#title      Isolation of human erythropoietin with monoclonal antibodies.
#cross-references M01D:84135751
#accession  A22210
#molecule_type protein
#residues   28-29,'X',31-33,'L',35-50,'X',52-53,'D',55,'G',57
            ##label YAN

COMMENT      Erythropoietin is produced by kidney or liver of adult mammals and
            by liver of fetal or neonatal mammals.

GENETICS
#gene       GDB:EPO
#map position 7q21.3-q22.1
#introns    5/1; 53/3; 82/3; 142/3
CLASSIFICATION
#superfamily erythropoietin
KEYWORDS     erythropoiesis; glycoprotein; hormone; kidney; liver
FEATURES
1-27        #domain signal sequence #status predicted #label SIGV
28-193      #product erythropoietin #status experimental #label MATV
34-188,56-60 #disulfide bonds #status experimental
51,65,110   #binding site carbohydrate (Aan) (covalent) #status
            experimentally
153         #binding site carbohydrate (Ser) (covalent) #status
            experimental
SUMMARY      #length 193 #molecular-weight 21307 #checksum 6371

Query Match      5.7%; Score 141; DB 2; Length 193;
Best Local Similarity 27.6%; Pred. No. 1.19e-03;
Matches 27; Conservative 27; Mismatches 41; Indels 3; Gaps 3;

Db 12 LILLELLEPLGLPVLGAPRLICDSRYLYLEAKAENITGCAHCELNITPTD 71
    ||| : ||| : | : | : || ||| : : : : | : | : | :
Oy 7 LLVYMLLTARLT-LSSPAPACDLRYLSKLIRDSHYLSRISQCPVHPLETPVLLPA 65

Db 72 kmfYawk-rmevvgqavevwqglalleaavlrqgall 108
    :: || ||| : | : : ||| ||| : : : | : | : | :
Oy 66 DFLSLGEMKTQME-ETKADIDILGAVTLLLEGVMAARGQL 102

RESULT      9
ENTRY      LURB11     #type complete
TITLE      annexin XI - rabbit
ALTERNATE_NAMES
ORGANISM    calyculin-associated annexin protein CAP-50
#formal_name Oryctolagus cuniculus #common_name domestic
            rabbit
DATE       30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change
24-Feb-1995

ACCESSIONS
JH0694; PH0950; A38250; PS0263
REFERENCE
JH0694
#authors    Tokumitsu, H.; Mizutani, A.; Muramatsu, M.; Yokota, T.; Arai,
            K.; Hidaka, H.

```

```

#journal      Biochem. Biophys. Res. Commun. (1992) 186:1227-1235
#title       Molecular cloning of rabbit CAP-50, a calyculin-associated
#            annexin protein.
#cross-references NID:92378579
#accession   JH0694
#molecule_type mRNA
#residues    1-503 ##label TOK
#cross-references DDBJ:D10883
#experimental source Lung
#accession   PH0550
#molecule_type protein
#residues    104-141;213-231;254-262;270-280;285-309;319-337;429-448;
            478-492 ##label PK2

REFERENCE
#authors     A38250
            Tokumitsu, H.; Mizutani, A.; Minami, H.; Kobayashi, R.;
            Hidaka, H.
#journal      J. Biol. Chem. (1992) 267:8919-8924
#title       A calyculin-associated protein is a newly identified member
            of the Ca(2+)/phospholipid-binding proteins, annexin
            family.
#cross-references NID:92250478
#accession   A38250
#molecule_type protein
#residues    104-141;213-223,'X',225-231;254-263;271-280;285-291,'X',
            293-300,'X',302-309;319-337;429-448;478-492 ##label
            TOK3

COMMENT      This protein binds specifically to calyculin in a Ca2+ dependent
            manner.

CLASSIFICATION #superfamily annexin VII; annexin repeat homology
KEYWORDS      calcium binding; duplication; endonexin fold; glycoprotein;
            phospholipid binding

FEATURE
201-272      #domain annexin repeat homology #label AX1\
212-228      #region endonexin fold #status predicted\
273-344      #domain annexin repeat homology #label AX2\
356-400      #region endonexin fold #status predicted\
356-428      #domain annexin repeat homology #label AX3\
366-384      #region endonexin fold #status predicted\
433-503      #domain annexin repeat homology #label AX4\
443-459      #region endonexin fold #status predicted\
58           #binding site carbohydrate (Asn) (covalent) #status
            predicted

SUMMARY      #length 503 #molecular-weight 54034 #checksum 4870

Query Match      5.6%; Score 139; DB 2; Length 503;
Best Local Similarity 25.7%; Pred. No. 1.94e-03;
Matches 26; Conservative 30; Mismatches 41; Indels 4; Gaps 4;

Db 36 glndvanyaqgfdyldsgmaamstfgnagvnpvlylgagp-gypppppgqfgqppp 94
::: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 239 SLDQJPGVIARI-HEILNGTGLGPEGPSRRRLGAPDISSTGDSLEPMLPGYS-PSP 296
::: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 95 tqpsvppvygyvpppggnppsgvpsypfpgapvpqgqmppp 135
::: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 297 THPPTGQYTLFPLPPLTPVIVQHLPLDDPSAVT-PPTTS 336
::: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 10
ENTRY   A33647
TITLE   sulfated surface glycoprotein 185 - Volvox carterii
ORGNISM #formal name Volvox carterii
DATE    11-Apr-1990 #sequence_revision 11-Apr-1990 #text_change
            31-Dec-1993
ACCESSION A33647

```

```
Bb      12 |||lsvlpqlpyppaprrllcdsrwlvrylleakeaaentvmgscssclnemi|typrdt   71  
||||::|| | ::::::|| ||||:: | : | : | : | : | : | : |  
  
Db      12 |||lsvlpqlpyppaprrllcdsrwlvrylleakeaaentvmgscssclnemi|typrdt   71  
||||::|| | ::::::|| ||||:: | : | : | : | : | : | : |
```

```
REFERENCE A33647  
#authors Ertl, H.; Mengede, R.; Wenzl, S.; Engel, J.; Stumper, M.  
#journal J Cell Biol. (1989) 109:3493-3501  
#title The extracellular matrix of Volvox carteri: molecular structure of the cellular compartment.
```

```
#cross-references NUID:90094551  
#accession A33647  
#status preliminary  
##molecule_type mRNA  
##residues 1-485 ##label ERT  
##cross-references GB:X51616  
KEYWORDS glycoprotein  
SUMMARY #length 485 #molecular-weight 50436 #checksum 3012
```

```
Query Match          5.3%; Score 132; DB 8; Length 485;  
Best Local Similarity 33.8%; Pred. No. 1.05e-02;  
Matches    24; Conservative    20; Mismatches 24; Indels     3; Gaps     3;
```

```
Dd      212 lsgnnvnpigqnamneblppepqgtlaasrpsppspprprrpppsps-psppppppppppp 270  
Qy       269 LGADPIS-G-TSDTGSLEPNIGVSPSTPPPTGQTLPFLPDTPPVQLHPLLD 326
```

```
Dd      271 ppppppsppp 281  
Qy       327 PSAPTPTPRSP 337
```

```
RESULT 11  
ENTRY erythropoietin precursor - crab-eating macaque  
TITLE erythropoietin precursor - crab-eating macaque  
ORGANISM Macaca fascicularis #common_name crab-eating macaque  
DATE 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 03-Mar-1995
```

```
ACCESSIONS J00173  
REFERENCE J00173  
#authors Lin, F.K.; Lin, C.H.; Lai, P.H.; Browne, J.K.; Egrie, J.C.; Smalling, R.; Fox, G.M.; Chen, K.K.; Castro, M.; Suggs, S.  
#journal Gene (1986) 44:201-209  
#title Monkey erythropoietin gene: cloning, expression and comparison with the human erythropoietin gene.  
#cross-references MUIID:87055236  
#accession U00173  
##molecule_type mRNA  
##residues 1-192 ##label LIN  
##experimental_source kidney  
COMMENT This protein is the principal hormone involved in the regulation of erythrocyte differentiation and the maintenance of a physiological level of circulating erythrocyte mass.  
CLASSIFICATION #superfamily erythropoietin  
KEYWORDS erythropoiesis; glycoprotein; hormone  
FEATURE  
1-24 #domain signal sequence #label SIG  
25-192 #product erythropoietin #label MATN  
51,65,110 #binding_site carbohydrate (asn) (covalent) #status predicted
```

```
SUMMARY #length 192 #molecular-weight 21113 #checksum 5284
```

```
Query Match          5.1%; Score 127; DB 5; Length 192;  
Best Local Similarity 25.3%; Pred. No. 3.4le-02;  
Matches    25; Conservative    30; Mismatches 41; Indels     3; Gaps     3;
```





The diagrams show the progression of the bubble sort algorithm:

- Diagram 1:** Initial array: 2, 10, 4, 7, 9, 3.
- Diagram 2:** After the first pass, the largest element (10) has bubbled to the end. Array: 2, 4, 7, 9, 3, 10.
- Diagram 3:** After the second pass, the next largest element (9) is in its final position. Array: 2, 4, 7, 3, 9, 10.
- Diagram 4:** After the third pass, the next largest element (7) is in its final position. Array: 2, 4, 3, 7, 9, 10.
- Diagram 5:** After the fourth pass, the next largest element (4) is in its final position. Array: 2, 3, 4, 7, 9, 10.
- Diagram 6:** The final sorted array is 2, 3, 4, 7, 9, 10.

(TM)

Release 2.1D John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.  
Distribution rights by IntelliGenetics, Inc.

MPsrch\_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Mon Apr 1 11:00:24 1996; MasPar time 7.11 Seconds

Tabular output not generated.

Title: >US-08-347-748-2

Item	Description:	Unit	Quantity	Unit Price	Total Price
1	(1-353) from US08347748.pdf		1	3403	3403

Sequence: 1 MELTELLVVMILTARTL.....PTSPINISYTHSQNSQEC 353

Scoring table: PAM 150

## Gap 11

Searched: 43470 seqs, 15335248 residues

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database: swiss-prot31

1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7  
8:part8

Statistics: Mean 48.744; Variance 109.173; scale 0.446

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result	No.	Score	Query Match	Length	DB	ID	Description	Pred. No.
1	2493	100.0	353	7	TPO_HUMAN		THROMBOPOIETIN PRECUR	0.00e+00
2	1748	70.1	356	7	TPO_MOUSE		THROMBOPOIETIN PRECUR	2.35e-294
3	162	6.5	192	3	EPO_MOUSE		ERYTHROPOIETIN PRECUR	3.81e-08
4	156	6.3	188	3	EPO_FELCA		ERYTHROPOIETIN PRECUR	2.58e-07
5	155	6.2	175	3	CMFCA		EYTHROPOIETIN PRECUR	3.55e-07
6	154	6.2	192	3	EPO_RAT		EYTHROPOIETIN PRECUR	4.84e-07
7	141	5.7	193	3	EPO_HUMAN		EYTHROPOIETIN PRECUR	2.67e-05
8	139	5.6	503	1	ANXA_RABIT		ANNEKIN XI (CALYCICLIN	4.88e-05
9	138	5.5	194	3	EPO_SHEEP		EYTHROPOIETIN PRECUR	6.59e-05
10	132	5.3	485	7	SSGG_VOLCA		SULFATED SURFACE GLYC	3.90e-04
11	127	5.1	192	3	EPO_MACFA		EYTHROPOIETIN PRECUR	1.66e-03
12	126	5.1	503	1	ANXA_BOVIN		ANNEKIN XI VARIANT 1	2.22e-03
13	116	4.7	751	8	VRP1_YEAST		VERPROLIN.	3.63e-02

15	14	13	12	11	10	9	8	7	6	5	4	3	2	1
45	44	43	42	41	40	39	38	37	36	35	34	33	32	31
28	27	26	25	24	23	22	21	20	19	18	17	16	15	14
15	14	13	12	11	10	9	8	7	6	5	4	3	2	1
45	44	43	42	41	40	39	38	37	36	35	34	33	32	31
28	27	26	25	24	23	22	21	20	19	18	17	16	15	14
15	14	13	12	11	10	9	8	7	6	5	4	3	2	1
45	44	43	42	41	40	39	38	37	36	35	34	33	32	31
28	27	26	25	24	23	22	21	20	19	18	17	16	15	14
15	14	13	12	11	10	9	8	7	6	5	4	3	2	1
45	44	43	42	41	40	39	38	37	36	35	34	33	32	31
28	27	26	25	24	23	22	21	20	19	18	17	16	15	14
15	14	13	12	11	10	9	8	7	6	5	4	3	2	1
45	44	43	42	41	40	39	38	37	36	35	34	33	32	31
28	27	26	25	24	23	22	21	20	19	18	17	16	15	14
15	14	13	12	11	10	9	8	7	6	5	4	3	2	1
45	44	43	42	41	40	39	38	37	36	35	34	33	32	31
28	27	26	25	24	23	22	21	20	19	18	17	16	15	14
15	14	13	12	11	10	9	8	7	6	5	4	3	2	1
45	44	43	42	41	40	39	38	37	36	35	34	33	32	31
28	27	26	25	24	23	22	21	20	19	18	17	16	15	14
15	14	13	12	11	10	9	8	7	6	5	4	3	2	1
45	44	43	42	41	40	39	38	37	36	35	34	33	32	31
28	27	26	25	24	23	22	21	20	19	18	17	16	15	14
15	14	13	12	11	10	9	8	7	6	5	4	3	2	1
45	44	43	42	41	40	39	38	37	36	35	34	33	32	31
28	27	26	25	24	23	22	21	20	19	18	17	16	15	14
15	14	13	12	11	10	9	8	7	6	5	4	3	2	1
45	44	43	42	41	40	39	38	37	36	35	34	33	32	31
28	27	26	25	24	23	22	21	20	19	18	17	16	15	14
15	14	13	12	11	10	9	8	7	6	5	4	3	2	1
45	44	43	42	41	40	39	38	37	36	35	34	33	32	31
28	27	26	25	24	23	22	21	20	19	18	17	16	15	14
15	14	13	12	11	10	9	8	7	6	5	4	3	2	1
45	44	43	42	41	40	39	38	37	36	35	34	33	32	31
28	27	26	25	24	23	22	21	20	19	18	17	16	15	14
15	14	13	12	11	10	9	8	7	6	5	4	3	2	1
45	44	43	42	41	40	39	38	37	36	35	34	33	32	31
28	27	26												

## ALIGNMENTS

ID	TPQ	HUMAN	STANDARD	PRT	353	AA.
AC	P40725					
DT	01-FEB-1995	(REL. 31, CREATED)				
DT	01-FEB-1995	(REL. 31, LAST SEQUENCE UPDATE)				
DT	01-FEB-1995	(REL. 31, LAST ANNOTATION UPDATE)				
DE	THROMBOPOIETIN PRECURSOR (MEGAKARYOCYTE COLONY STIMULATING FACTOR)					
DE	(C-MPL LIGAND) (ML) (MEGAKARYOCYTE GROWTH AND DEVELOPMENT FACTOR)					
GN	(MGPF) .					
OS	THPO.					
OS	HOMO SAPIENS (HUMAN) .					
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;					
OC	EUTHERIA; PRIMATES.					
LN	(1)					
RP	SEQUENCE FROM N.A.					
RM	94261202					
RA	DE SAUVAGE F.J., HASS P.E., SPENCER S.D., MALLOY B.E., GUREY A.L.,					
RA	SPENCER S.A., DARBOONE W.C., HENZEL W.J., WONG S.C., KUANG M.-J.,					
RA	OLDS K.J., HULTGREEN B., SOLBERG L.A. JR., GOEDDEL D.V., EATON D.L.,					
RL	NATURE 369:533-538(1994).					
LN	(2)					
RP	SEQUENCE FROM N.A.					
RM	94291201					
RC	TISSUE=LIVER;					
RA	BARTLEY T.D., BOGENBERGER J., HUNT P., LI Y.-S., LU H.-S.,					
RA	MARTIN F., CHANG M.-S., SAMAL B.B., NICHOL S., BOSSELMAN R.A.;					
RL	CELL 77:1117-1124(1994).					
CC	-I- FUNCTION: ACTS AS A CIRCULATING REGULATOR OF PLATELET NUMBERS.					

CC MEGAKARYOCYTIC LINEAGE SPECIFIC GROWTH AND DIFFERENTIATION  
 CC FACTOR.  
 CC -1- SUBCELLULAR LOCATION: SECRETED.  
 DR EMBL; L33410; HSMCLMPL.  
 DR EMBL; U11025; HS11025.  
 DR PIR; S45331; S45331.  
 DR MIM; 600044; 11TH EDITION.  
 KM GLYCOPROTEIN; HORMONE; SIGNAL.  
 FT SIGNAL 1 21 POTENTIAL.  
 FT CHAIN 22 353 THROMBOPOIETIN.  
 FT CARBOHYD 197 197 POTENTIAL.  
 FT CARBOHYD 206 206 POTENTIAL.  
 FT CARBOHYD 234 234 POTENTIAL.  
 FT CARBOHYD 255 255 POTENTIAL.  
 FT CARBOHYD 340 340 POTENTIAL.  
 FT CARBOHYD 348 348 POTENTIAL.  
 FT CONFLICT 113 113 Q -> E (IN REF. 2).  
 SQ SEQUENCE 353 AA; 37822 MW; 702923 CN;

Query Match 100.0%; Score 2493; DB 7; Length 353;  
 Best Local Similarity 100.0%; Pred. No. 0.00e+00;  
 Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 melteellllyvmltlcarltlesppapadrlvleklldshvherlegcpvhpilpvp 60  
 Qy 1 MELTEELLVYMLTLARLTLSPPACDRLVLSKLSDSHVLSHLSQCEVHPILPVP 60  
 Db 61 lrpavdfelegwktqmeekadilgavrtlllegvmaargqlpvtclselllgqlsgvrl 120  
 Qy 61 LRPADVFELEGWKTQMEETKADILGAVRTLLLEGVMAARGQLPVTCLSELGQLSGVRL 120  
 Db 121 llygalqslgtcqlpqgrttahkdpnaiflsfqhlrtgkvrfilmvgseclvrrisapt 180  
 Qy 121 LLYGALQSLGTCLPQGRTTAHKDPNAIFLSFQHLRTGKVRFLMVGSECLVRRISAPT 180  
 Db 121 llygalqslgtcqlpqgrttahkdpnaiflsfqhlrtgkvrfilmvgseclvrrisapt 180  
 Qy 121 LLYGALQSLGTCLPQGRTTAHKDPNAIFLSFQHLRTGKVRFLMVGSECLVRRISAPT 180  
 Db 181 avpsatqltlcinkfntegqlletnfsvartacagpqllerlgfvtktpqqlnqere 240  
 Qy 181 AVPSATSLVLTINELPNTSGLENTFTASARTTSGGLKMQGFRAKIPGLANQTSNSL 240  
 Db 241 dqipjmrthlbntrgqlfpgprertlgaipdisetdsqglpnqlqpyapapsthp 300  
 Qy 241 DQIPGTYMRHIELNCTRGLEPGPSRRTLGAPDISSETDSGLPNQLQPYAPAPSTHP 300  
 Db 301 tgyvclfrpplclprrvqqlhpllpdpapcpptcepllnlsvtchmqleaqeg 353  
 Qy 301 TGQVCLFRPPLCLPRRVQQLHPLLPDPAPCPPTCEPLNLSTVCHMSQLEAQEG 353

RESULT 2  
 ID TPO MOUSE STANDARD; PRT; 356 AA.  
 AC P40226;  
 DT 01-FEB-1995 (REL. 31, CREATED)  
 DT 01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)  
 DT 01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)  
 DE THROMBOPOIETIN PRECURSOR (MEGAKARYOCYTE COLONY STIMULATING FACTOR)  
 DE (C-MPL LIGAND) (ML) (MEGAKARYOCYTE GROWTH AND DEVELOPMENT FACTOR)  
 DE (MGDF).  
 OS MUS MUSCULUS (MOUSE).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;  
 GN EUTHERIA; RODENTIA.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RM 94261207  
 RA LOK S., KAUSHANSKY K., HOLLY R.D., KUTJER J.L., LOFTON-DAY C.E.,

RA OORT P.J., GRANT F.J., HEIPEL M.D., BURKHEAD S.K., KRAMER J.M.,  
 RA BELL L.A.N., SPEICHER C.A., BLUMBERG H., JOHNSON R., PRONKARD D.,  
 RA CHING A.F.T., MATHEWS S.L., BAILEY M.C., FORSTROM J.M., BUDDLE M.M.,  
 RA OSBORNE S.G., EVANS S.J., SHEPPARD P.O., PRESNELL S.R., O'HARA P.J.,  
 RA HAGEN F.S., ROTH G.J., FOSTER D.C.;  
 RL NATURE 369:565-568 (1994).  
 CC -1- FUNCTION: ACTS AS A CIRCULATING REGULATOR OF PLATELET NUMBERS.  
 CC -1- MEGAKARYOCYTIC LINEAGE SPECIFIC GROWTH AND DIFFERENTIATION  
 CC FACTOR.  
 CC -1- SUBCELLULAR LOCATION: SECRETED.  
 DR EMBL; L34169; MATHROA.  
 DR PIR; S45330; S45330.  
 KM GLYCOPROTEIN; HORMONE; SIGNAL.  
 FT SIGNAL 1 21 POTENTIAL.  
 FT CHAIN 22 356 THROMBOPOIETIN.  
 FT CARBOHYD 197 197 POTENTIAL.  
 FT CARBOHYD 206 206 POTENTIAL.  
 FT CARBOHYD 235 235 POTENTIAL.  
 FT CARBOHYD 249 249 POTENTIAL.  
 FT CARBOHYD 256 256 POTENTIAL.  
 FT CARBOHYD 336 336 POTENTIAL.  
 FT CARBOHYD 351 351 POTENTIAL.  
 SQ SEQUENCE 356 AA; 37835 MW; 701294 CN;

Query Match 70.1%; Score 1748; DB 7; Length 356;  
 Best Local Similarity 71.6%; Pred. No. 2.32e-294;  
 Matches 255; Conservative 47; Mismatches 49; Indels 5; Gaps 4;

Db 1 melteellllyvmltlcarltlesppapadrlvleklldshvherlegcpvhpilpvp 60  
 Qy 1 MELTEELLVYMLTLARLTLSPPACDRLVLSKLSDSHVLSHLSQCEVHPILPVP 60  
 Db 61 lrpavdfelegwktqmeekadilgavrtlllegvmaargqlpvtclselllgqlsgvrl 120  
 Qy 61 LRPADVFELEGWKTQMEETKADILGAVRTLLLEGVMAARGQLPVTCLSELGQLSGVRL 120  
 Db 121 llygalqslgtcqlpqgrttahkdpnaiflsfqhlrtgkvrfilmvgseclvrrisapt 180  
 Qy 121 LLYGALQSLGTCLPQGRTTAHKDPNAIFLSFQHLRTGKVRFLMVGSECLVRRISAPT 180  
 Db 181 avpsatqltlcinkfntegqlletnfsvartacagpqllerlgfvtktpqqlnqere 240  
 Qy 181 AVPSATSLVLTINELPNTSGLENTFTASARTTSGGLKMQGFRAKIPGLANQTSNSL 240  
 Db 241 dqipjmrthlbntrgqlfpgprertlgaipdisetdsqglpnqlqpyapapsthp 300  
 Qy 241 DQIPGTYMRHIELNCTRGLEPGPSRRTLGAPDISSETDSGLPNQLQPYAPAPSTHP 300  
 Db 301 tgyvclfrpplclprrvqqlhpllpdpapcpptcepllnlsvtchmqleaqeg 355  
 Qy 301 TGQVCLFRPPLCLPRRVQQLHPLLPDPAPCPPTCEPLNLSTVCHMSQLEAQEG 355

RESULT 3  
 ID EPO MOUSE STANDARD; PRT; 192 AA.  
 AC P07321;  
 DT 01-APR-1988 (REL. 07, CREATED)  
 DT 01-APR-1988 (REL. 07, LAST SEQUENCE UPDATE)  
 DT 01-FEB-1994 (REL. 28, LAST ANNOTATION UPDATE)  
 DE ERYTHROPOIETIN PRECURSOR.  
 GN EPO.  
 OS MUS MUSCULUS (MOUSE).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;  
 GN EUTHERIA; RODENTIA.



RN	[1]	SEQUENCE FROM N.A.	
RP	87039105		
RA	SHOEMAKER C.B., MITSOCK L.D.;		
RL	MOL. CELL. BIOL. 6:849-858(1986).		
RN	(2)		
RP	SEQUENCE FROM N.A.		
RM	87039104		
RA	MC DONALD J.D., LIN F.-K., GOLDMANSER E.;		
RL	MOL. CELL. BIOL. 6:842-848(1986).		
CC	- - FUNCTION: ERYTHROPOIETIN IS THE PRINCIPAL HORMONE INVOLVED IN THE REGULATION OF ERYTHROCYTE DIFFERENTIATION AND THE MAINTENANCE OF A PHYSIOLOGICAL LEVEL OF CIRCULATING ERYTHROCYTE MASS.		
CC	- - TISSUE SPECIFICITY: PRODUCED BY KIDNEY OR LIVER OF ADULT MAMMALS AND BY LIVER OF FETAL OR NEONATAL MAMMALS.		
CC	- - SUBCELLULAR LOCATION: SECRETED.		
DR	EMBL; M12482; MMERP.		
DR	PIR; A24901; A24901.		
DR	PIR; A24902; A24902.		
DR	PROSITE; PS00817; EPO.		
KW	ERYTHROCYTE MATURATION; GLYCOPROTEIN; HORMONE; SIGNAL.		
FT	SIGNAL 1 26		
FT	CHAIN 27 192	ERYTHROPOIETIN.	
FT	DISULFID 33 187	BY SIMILARITY.	
FT	CARBOHYD 50 50	BY SIMILARITY.	
FT	CARBOHYD 64 64	BY SIMILARITY.	
FT	CARBOHYD 109 109	BY SIMILARITY.	
SQ	SEQUENCE 192 AA; 21365 MW; 188541 CN;		
 Query Match 6.5%; Score 162; DB 3; Length 192; Best Local Similarity 27.6%; Pred. No. 3,8le-08; Matches 35; Conservative 34; Mismatches 54; Indels 4; Gaps 4;			
Db	10 [     ]lpvlpcaprilcdsrlyrlylkaeaenvtmgcaegprlsenttpd 69 :::      -   -:         : : : :     : : :		
Qy	6 LLIVVALLITLARTLSSPAPACDLAVLSKLRLDSHVYLSRLSQCPVEHPLRPVYLDA 64		
Db	70 tkynfwawk-tmevegeaeavwqglaleallseqallansgppetqlhdksaisgl 128 ::   :       : ::      :   : : : :    : :		
Qy	65 VDSISLEKTKTQME-ETKADDIIGANTLLIEGVMAARGQLGPTCLSLSLQSLSGVRL-L 122		
Db	129 rrltsll 135 :		
Qy	123 GALDSL 129		
 RESULT 4 ID EPO FELCA STANDARD; PRT; 188 AA. AC P33708; DT 01-FEB-1994 (REL. 28, CREATED) DT 01-FEB-1994 (REL. 28, LAST SEQUENCE UPDATE) DT 01-FEB-1994 (REL. 28, LAST ANNOTATION UPDATE) DE ERYTHROPOIETIN PRECURSOR (FRAGMENT). GN EPO. OS FELIS CATUS (CAT). OC EUMARVORA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA; OC EUTHERIA; CARNIVORA. RN [1] RP SEQUENCE FROM N.A. RM 93372347 RA WEN D., BOISSEL J.P.R., TRACY T.E., MOLCAHY I.S., CZELUSNIAK J., RA GOODMAN M., BUNN H.F.; RL BLOOD 82:1507-1516(1993).			

CC	-1- FUNCTION: ERYTHROPOIETIN IS THE PRINCIPAL HORMONE INVOLVED IN THE REGULATION OF ERYTHROCYTE DIFFERENTIATION AND THE MAINTENANCE OF A PHYSIOLOGICAL LEVEL OF CIRCULATING ERYTHROCYTE MASS.
CC	-1- TISSUE SPECIFICITY: PRODUCED BY KIDNEY OR LIVER OF ADULT MAMMALS AND BY LIVER OF FETAL OR NEONATAL MAMMALS.
CC	-1- SUBCELLULAR LOCATION: SECRETED.
DR	EMBL; L10606; FCERYTHRO.
DR	PROSITE; PS00817; EPO.
KM	ERYTHROCYTE MATURATION; GLYCOPROTEIN; HORMONE; SIGNAL.
FT	NON TER 1 1
FT	SIGNAL <1 22 BY SIMILARITY.
FT	CHAIN 23 188 ERYTHROPOIETIN.
FT	DISULEID 29 183 BY SIMILARITY.
FT	DISULEID 51 55 BY SIMILARITY.
FT	CARBOHYD 46 46 POTENTIAL.
FT	CARBOHYD 60 60 POTENTIAL.
FT	CARBOHYD 105 105 POTENTIAL.
SQ	SEQUENCE 188 AA; 20608 MM; 175419 CN;
Query Match	6.3%; Score 156; DB 3; Length 188;
Best Local Similarity	24.8%; Pred. No. 2,58e-07;
Matches	36; Conservative 45; Mismatches 59; Indels 5; Gaps 5;
Db	6 LLLILLIIPILPIVLPQAPPLICDSFVLEYLLEAREAEVYMGCAEGCEAFENTYPD 65
Qy	6 LLLVWMLLTARLT-LSPPAPACDLSEVSKLLRDSHVLASRLSQCPEVHPDPTVLLPA 64
Db	66 tkmfvfk-tpndqgqavewqglalsallrgqalansgspeelqdhvksael 124
Qy	65 VDFSLGEMTKQWE-ETKADDIIGAVTLLLEGWAMRQGLPTCLSSLLGSLQSVRL-L 122
Db	125 relteilla-lgaqkeateapeate 148
Qy	123 GALQSLIGTQLPFGGRTTHAKDPNA 147
RESULT	5
ID	EPO CANFA STANDARD; PRT; 175 AA.
AC	P33707;
DT	01-FEB-1994 (REL. 28, CREATED)
DT	01-FEB-1994 (REL. 28, LAST SEQUENCE UPDATE)
DT	01-FEB-1994 (REL. 28, LAST ANNOTATION UPDATE)
DE	ERYTHROPOIETIN PRECURSOR (FRAGMENT).
GN	EPO.
OS	CANIS FAMILIARIS (DOG).
OS	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC	EUTHERIA; CARNIVORA.
RN	(1)
RP	SEQUENCE FROM N.A.
RM	93372347
RA	WEN D., BOISSEL J.P.R., TRACY T.E., MULCAHY L.S., CZELUSNIAN J.,
RA	GOODMAN M., BUNN H.F.;
RL	BLOOD 82:1507-1516(1993).
CC	-1- FUNCTION: ERYTHROPOIETIN IS THE PRINCIPAL HORMONE INVOLVED IN THE REGULATION OF ERYTHROCYTE DIFFERENTIATION AND THE MAINTENANCE OF A PHYSIOLOGICAL LEVEL OF CIRCULATING ERYTHROCYTE MASS.
CC	-1- TISSUE SPECIFICITY: PRODUCED BY KIDNEY OR LIVER OF ADULT MAMMALS AND BY LIVER OF FETAL OR NEONATAL MAMMALS.
CC	-1- SUBCELLULAR LOCATION: SECRETED.
DR	EMBL; L13021; FCERYPRE.
DR	PROSITE; PS00817; EPO.
KM	ERYTHROCYTE MATURATION; GLYCOPROTEIN; HORMONE; SIGNAL.
FT	NON TER 1 1
FT	SIGNAL <1 22 BY SIMILARITY.



RH		STRUCTURE OF CARBOHYDRATES.
RM	92314463	
RA	TAKEUCHI M., KOBATA A.;	
RL	GLYCOBIOLOGY 1:337-346(1991).	
RC	-!- FUNCTION: ERYTHROPOIETIN IS THE PRINCIPAL HORMONE INVOLVED IN THE	
RD	REGULATION OF ERYTHROCYTE DIFFERENTIATION AND THE MAINTENANCE OF A	
RE	PHYSIOLOGICAL LEVEL OF CIRCULATING ERYTHROCYTE MASS.	
RF	-!- TISSUE SPECIFICITY: PRODUCED BY KIDNEY OR LIVER OF ADULT MAMMALS	
RG	AND BY LIVER OF FETAL OR NEONATAL MAMMALS.	
RH	-!- SUBCELLULAR LOCATION: SECRETED.	
RI	EMBL; X02158; HSERP.G.	
RJ	EMBL; X02157; HSERP.R.	
RK	EMBL; M11319; HSERP.A.	
RL	PIR; A01855; ZOHU.D.	
RM	PIR; A25384; A25384.	
RN	PIR; A24744; A24744.	
RP	PIR; A22210; A22210.	
RQ	MIM; I33170; 11TH EDITION.	
RS	PROSITE; PS00817; EPO.	
RT	KW ERYTHROCYTE MATURATION; GLYCOPROTEIN; HORMONE; SIGNAL.	
RU	FT SIGNAL 1 27	
RV	FT CHAIN 28 193	ERYTHROPOIETIN.
RW	FT DISULEFD 34 188	
RX	FT DISULEFD 56 60	
RY	FT CARBOHYD 51 51	
RZ	FT CARBOHYD 65 65	
S0	FT CARBOHYD 110 110	
S1	FT CARBOHYD 153 153	
S2	FT PROPEP 190 193	
S3	FT CONFLICT 85 85	MAY BE REMOVED IN PROCESSED PROTEIN. 0 -> 00 (IN REF. 3).
S4	SEQUENCE 193 AA; 21306 MW; 182311 CN;	
Dc	Query Match 5.7%; Score 141; DB 3; Length 193; Best Local Similarity 27.6%; Pred. No. 2.67e-05; Matches 27; Conservative 27; Mismatches 41; Indels 3; Gaps 3	
Dd	Db 12      lelpldlpylgpprlcdervlylleakeenttgcacnclsnhticvpt 71 ::   :         :   :   :   :   :   :   :   :   :   : 1 LVAVMLLLTALLTL--LSSPAAPACDIARVLSTLRDSHVLAHSLSGCEVHPLEPTVLPAAV 65	
Df	72 kmfyawk-tmevgqgavevwggllalalseevlrngall 108 ::   :         : : : :   :   :   :   :   :   :   : 66 DFLSIEGMTQMGE-ETRKAODILCAVTLVILLECVMAARGQL 102	
RESULT	8 STANDARD; PRT; 503 AA.	
ID ANYA RABIT	P33477;	
DT 01-FEB-1994 (REL. 28, CREATED)		
DT 01-FEB-1994 (REL. 28, LAST SEQUENCE UPDATE)		
DE ANNEXIN XI (CALCYLEGIN-ASSOCIATED ANNEXIN 50) (CAP-50),		
OS OXYCTOLAAGUS CINICTULUS (RABBIT).		
OC EUCHARIOTA; METAZOIA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;		
CX EUTHERIA; LACOMORPHA.		
CC [1]		
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.		
RC TISSUE=LUNG;		
RM 92378579		
RA TOKUMITSU H., MIIZUTANI A., MORAHATSU M., YOKOTA T., ARAI K.,		
HIDAHA H.;		
BIOCHEM. BIOPHYS. RES. COMMUN. 186:1227-1235(1992).		
-!- FUNCTION: BINDS SPECIFICALLY TO CALCYLEGIN IN A CALCIUM-DEPENDENT MANNER.		

```
CC      -1- DOMAIN: CONTAINS FOUR HOMOLOGOUS REPEATS WITH A CONSENSUS  
CC      SEQUENCE COMMON TO ALL ANNEXIN PROTEINS. A PAIR OF THESE REPEATS  
CC      MAY FORM ONE BINDING SITE FOR CALCIUM AND PHOSPHOLIPID.  
CC      -1- SIMILARITY: TO OTHER PROTEINS OF THE ANNEXIN FAMILY.  
DR      EMBL, D10883; OCCAP.  
DR      PIR, JH0694; IJRB11.  
DR      HSPD, P19999; ICLE.  
DR      PROSITE, PS00223; ANNEXIN.  
KM      ANNEXIN; CALCIUM/PHOSPHOLIPID-BINDING; REPEAT.  
FT      REPEAT       207   267     ANNEXIN.  
FT      REPEAT       279   339     ANNEXIN.  
FT      REPEAT       363   423     ANNEXIN.  
FT      REPEAT       438   498     ANNEXIN.  
SQ      SEQUENCE    503 AA; 54034 MW; 1231578 CN;  
  
Query Match          5.6%; Score 139; DB 1; Length 503;  
Best Local Similarity 25.7%; Pred. No. 4,88e-05;  
Matches 26; Conservative 30; Mismatches 41; Indels 4; Gaps 4;  
  
Db      36 glshvayagdfngdyLegmaamqgffgganvpnilypqag-gyvpvyqgfqqpp 94  
QY      239 SLIDQLPGYLKRI-HELLNGTRGLGFGPSRRRTLCADVDISSGTDSPLPVLQPGYS--PSP 296  
QY      297 THPPGTGYTLFPIPELTLPPTVTVQLPILPLPSAPR--PTPRS 336  
QY      297 THPPGTGYTLFPIPELTLPPTVTVQLPILPLPSAPR--PTPRS 336  
  
RESULT 9  
ID      EPO SHEEP           STANDARD;             PRT;         194 AA.  
AC      P33709;  
DT      01-FEB-1994 (REL. 28, CREATED)  
DT      01-FEB-1994 (REL. 28, LAST SEQUENCE UPDATE)  
DT      01-FEB-1994 (REL. 28, LAST ANNOTATION UPDATE)  
DE      ERYTHROPOIETIN PRECURSOR.  
GN      EPO.  
OS      OVIS ARIES (SHEEP).  
OC      EDUARDOYA, METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;  
OC      EUTHERIA; ARTIODACTYLIA.  
NC      [1]  
RP      SEQUENCE FROM N.A.  
RC      TISSUE=KIDNEY;  
RM      93351736  
  
RA      FU P., EVANS B., LIM G.B., MORITZ K., WINTOUR M.E.;  
RL      MOL. CELL. ENDOCRINOLOG., 93:107-116(1993).  
CC      -1- FUNCTION: ERYTHROPOIETIN IS THE PRINCIPAL HORMONE INVOLVED IN THE  
CC      REGULATION OF ERYTHROCYTE DIFFERENTIATION AND THE MAINTENANCE OF A  
CC      PHYSIOLOGICAL LEVEL OF CIRCULATING ERYTHROCYTE MASS.  
CC      -1- TISSUE SPECIFICITY: PRODUCED BY KIDNEY OR LIVER OF ADULT MAMMALS  
CC      AND BY LIVER OF FETAL OR NEONATAL MAMMALS.  
CC      -1- SUBCELLULAR LOCATION: SECRETED.  
DR      EMBL, Z24681; OAERYPOIA.  
DR      PROSITE, PS00817; EPO.  
KM      ERYTHROCYTE MATURATION; GLYCOPROTEIN; HORMONE; SIGNAL.  
FT      FT SIGNAL      1       27     BY SIMILARITY.  
FT      CHAIN          28   184     ERYTHROPOIETIN.  
FT      DISULFID       34   189     BY SIMILARITY.  
FT      DISULFID       56   60      BY SIMILARITY.  
FT      CARBOHYD       51   51      POTENTIAL.  
FT      CARBOHYD       65   65      POTENTIAL.  
FT      CARBOHYD      110   110     POTENTIAL.  
SQ      SEQUENCE    194 AA; 21335 MW; 175348 CN;
```

Dd	11	f  fpjg vpld april c sr lry leare aenatmgcaegcsfsen typd	70
Qy	6	LLLVWVILITLRLT-LSSPAPACDLYLSKLRDSHVHSLRSLSQCEVHPPLPVLLPA	64
Dd	71	tknfyawk-rmevgqga ewqgall sealirg galla	109
Qy	65	VDFSEIGEMKTQWE-ETKAODIILGAVTTLLLEGVAARGLG	103
 RESULT 10			
ID	SSGP VOLCA	STANDARD;	PRT; 485 AA.
AC	P2197;		
DT	01-AUG-1991 (REL. 19,	CREATED)	
DT	01-AUG-1991 (REL. 19,	LAST SEQUENCE UPDATE)	
DT	01-OCT-1994 (REL. 30,	LAST ANNOTATION UPDATE)	
DE	SULFATED SURFACE GLYCOPROTEIN 185 (SSG 185).		
OS	VOLVOX CARTERI.		
OC	EUKARYOTA; PLANTA; PHYCOPHYTA; CHLOROPHYTA (GREEN ALGAE);		
OC	CHLOROPHYCEAE; VOLVOCALES; VOLVOACEAE.		
RN	[1]		
RP	SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.		
RC	STRAIN=F. NAGARIENSIS / HK 10;		
RM	90094551		
RA	ERTL H.; MENGELE R.; WENZL S.; ENGEL J.; SUMPER M.;		
RL	J. CELL. BIOL. 109:3493-3501(1989).		
CC	- - FUNCTION: THE EXTRACELLULAR MATRIX (ECM) OF VOLVOX CONTAINS INSOLUBLE FIBROUS LAYERS THAT SURROUND INDIVIDUAL CELLS AT A DISTANCE TO FORM CONTIGUOUS CELLULAR COMPARTMENTS. SSG 185 IS THE MONOMERIC PRECURSOR OF THIS SUBSTRUCTURE (C32 STRUCTURE). THE COVALENT CROSS-LINKS ARE FORMED BETWEEN THE SACCHARIDE CHAINS. RATHER THAN BETWEEN THE POLYPEPTIDE CHAINS.		
CC	- - PTM: A CENTRAL DOMAIN OF 77 AA CONSISTS ALMOST EXCLUSIVELY OF HYDROXYPROLINE RESIDUES.		
DR	PIR; A33647; A33647.		
DR	HSSP; P19999; ICLG.		
KM	GLYCOPROTEIN; SULFATATION; HYDROXYLATION.		
FT	DOMAIN 228 340 PRO-RICH.		
FT	DOMAIN 260 295 POLY-PRO.		
SO	SEQUENCE 485 AA; 50436 MW; 1322558 CN;		
 Query Match 5.3%; Score 132; DB 7; Length 485; Best Local Similarity 33.8%; Pred. No. 3,90e-04; Matches 24; Conservative 20; Mismatches 24; Indels 3; Gaps 3;			
Dd	212	legpnvnpigapnnsplppspsaaesrppspssprpsppps-papppppppppppp	270
Qy	269	LGRADISS-G-TSDTSGLSPRNIGRGPSPTHTPRTCQTTLPELPPLTPVPVQLHLPLRD	326
Dd	271	pppppppppp 281	
Qy	327	PSAPTPTPTSP 337	
 RESULT 11			
ID	EPO MACFA	STANDARD;	PRT; 192 AA.
AC	P0786;		
DT	01-AUG-1988 (REL. 08,	CREATED)	
DT	01-AUG-1988 (REL. 08,	LAST SEQUENCE UPDATE)	
DT	01-FEB-1994 (REL. 28,	LAST ANNOTATION UPDATE)	
EN	ERYTHROPOIETIN PRECURSOR.		
EN	EPO.		

OS	MACACA FASCICULATA (CRAB EATING MACAQUE) (CYNOMOLGUS MONKEY)
OC	EUTHERIA, METAZOA, CHORDATA, VERTEBRATA, TETRAPODA, MAMMALIA,
OC	EUTHERIA, PRIMATES.
RM	(1)
RP	SEQUENCE FROM N.A.
RM	87055236
RA	LIN F.-K., LIN C.-H., LAI P.-H., BROWNE J.K., EGRIE J.C., SMALLING R.,
RA	FOX G.M., CHEN K.K., CASTRO M., SUGGS S.;
RL	GENE 44:201-209(1986).
CC	-1- FUNCTION: ERYTHROPOIETIN IS THE PRINCIPAL HORMONE INVOLVED IN THE
CC	REGULATION OF ERYTHROCYTE DIFFERENTIATION AND THE MAINTENANCE OF A
CC	PHYSIOLOGICAL LEVEL OF CIRCULATING ERYTHROCYTE MASS.
CC	-1- TISSUE SPECIFICITY: PRODUCED BY KIDNEY OR LIVER OF ADULT MAMMALS
CC	AND BY LIVER OF FETAL OR NEONATAL MAMMALS.
CC	-1- SUBCELLULAR LOCATION: SECRETED.
DR	EMBL; M18189; MFEPO.
DR	EMBL; M18188; MFEPO.
DR	PIR; J00173; J00173.
DR	PROSITE; PS00817; EPO.
KM	ERYTHROCYTE MATURATION; GLYCOPROTEIN; HORMONE; SIGNAL.
FT	SIGNAL 1 27 BY SIMILARITY.
FT	CHAIN 28 192 ERYTHROPOIETIN.
FT	DISULEID 34 187 BY SIMILARITY.
FT	DISULEID 56 60 BY SIMILARITY.
FT	CARBOHYD 51 51 BY SIMILARITY.
FT	CARBOHYD 65 65 BY SIMILARITY.
FT	CARBOHYD 110 110 BY SIMILARITY.
FT	CARBOHYD 152 152 BY SIMILARITY.
FT	SEQUENCE 192 AA; 21113 MW; 175216 CN;
Qy	Query Match 5.1%; Score 127; DB 3; Length 192;
Qy	Best Local Similarity 25.3%; Pred.No. 1.66e-03;
Qy	Matches 25; Conservative 30; Mismatches 41; Indels 3; Gaps 3;
Db	12 lllalvslplq pvpqaprlldcslrvlyllleakeavntmgsesce nenltvpt 71
Qy	7 LTVMLLLTARLTITSSPAP-ACDLRLVLSKLRDSHVLSRSLQCEPVEHPIPLPVLLAV 65
Db	72 kmfyawk-rmevqgqvevwgjal leaeavlrgavla 109
Qy	66 DFLSGMKTOME-ETRAQDILGAVTLLEGVMAARGLG 103
RESULT 12	
ID	ANYA BOVIN STANDARD; PRT; 503 AA.
AC	P27214;
DT	01-AUG-1992 (REL. 23, CREATED)
DT	01-AUG-1992 (REL. 23, LAST SEQUENCE UPDATE)
DT	01-OCT-1994 (REL. 30, LAST ANNOTATION UPDATE)
DE	ANNEXIN XI VARIANT 1 (CALCYCLIN-ASSOCIATED ANNEXIN 50) (CAP-50).
OS	BOS TAURUS (BOVINE).
OC	EUTHERIA, METAZOA, CHORDATA, VERTEBRATA, TETRAPODA, MAMMALIA,
OC	EUTHERIA, ARTIODACTYLIA.
RM	(1)
RP	SEQUENCE FROM N.A.
RC	TISSUE=CHONDROCYTE;
RM	92184796
RA	TOWLE C.A., TREADWELL B.V.;
RL	J. BIOL. CHEM. 267:5416-5423(1992).
RP	(2)
RC	SEQUENCE FROM N.A.
RC	TISSUE=CARTILAGE;
RM	92305067
RA	TOWLE C.A., WEISSBACH L., TREADWELL B.V.;

Dl	37	gldhvanyagdfngdyfsgyaamqtffgga-nvplypgapq-vgvpppgfsgppp	94
Oy	239	SIDDPETYLARI-HELLINGRGLFGPGSRRTLCAPDISSGTDSGLPNLQPEtS-PSP	296
Dd	95	aadpyepgympggprptcgsmps/ppypgpagpqpqqlpp	135
Oy	297	THPPGQTFLFPLPPTLPRTLVVQLRLPLPSAPPRPTSP	337
RESULT	13		
ID	VREP1 YEAST	STANDARD;	PRT; 751 AA.
AC	P37370;		
DT	01-OCT-1994 (REL. 30, CREATED)		
DT	01-OCT-1994 (REL. 30, LAST SEQUENCE UPDATE)		
DT	01-OCT-1994 (REL. 30, LAST ANNOTATION UPDATE)		
DE	VERPROLIN.		
GN	VREP1.		
OS	SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).		
OC	EUKARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOMETES.		
RN	(1)		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=A364;		
RH	95058201		
RA	DONNELLY S.F.H., POCKLINGTON M.J., PALLOTA D., ORR E.;		
RL	MOL. MICROBIOL. 10:585-596(1993).		
CC	-I- FUNCTION: INVOLVED IN CYTOSKELETAL ORGANISATION AND CELLULAR GROWTH. MAY EXERT ITS EFFECTS ON THE CYTOSKELETON DIRECTLY, OR INDIRECTLY VIA PROLINE-BINDING PROTEINS (E.G. PROFILIN) OR PROTEINS POSSESSING SH3 DOMAINS.		
CC	EMBL; Z26645; SVEPERRL.		
DR	PIR; S39626; S39626.		
KM	REPEAT.		
FT	DOMAIN	5	14 POLY-PRO.
FT	DOMAIN	239	245 POLY-PRO.
FT	DOMAIN	349	357 POLY-PRO.
FT	DOMAIN	396	406 POLY-PRO.
FT	DOMAIN	424	431 POLY-PRO.
FT	DOMAIN	462	468 POLY-SER.
FT	DOMAIN	704	708 POLY-PRO.
SO	SEQUENCE	751 AA;	75953 MW; 3267787 CN;

```

Query Match          4.7%; Score 116; DB 8; Length 751;
Best Local Similarity 28.4%; Pred. No. 3,63e-02;
Matches 19; Conservative 23; Mismatches 23; Indels 2; Gaps 2;

Db      374 pcdlapp|ntcvsyppn-kasempapppppppgafcsesalsaeaiplapl-ppppp 431
       :::: ::|||:: :|::| | |::|::|::| | |::|
Qy      271 ADLISGSDTGSPLPNLQGCVSPFTHPTGGYTLPLPPTLPVTVQLHLLDPSSA 330

Db      432 svatvcp 438
       : : : |
Qy      331 TPTPTSP 337

RESULT 14
ID YKF9 YEAST STANDARD; PRT; 441 AA.
AC P35728;
DT 01-JUN-1994 (REL. 29, CREATED)
DT 01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)
DE 01-OCT-1994 (REL. 30, LAST ANNOTATION UPDATE)
DR HYPOTHETICAL_49.6_KD PROTEIN IN FBA1-T0A2 INTERGENIC REGION.
GN YKL059C OR YKL316.
OS SACCAROMYCES CEREVISIAE (BAKER'S YEAST).
OC EUKARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOMETES.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C;
RM 94378723
RA RASMUSSEN S.W.;
RL YEAST 10:S63-S68(1994).
DR EMBL; X75781; SCX1286K.
DR EMBL; Z28059; SCTK1050C.
DR PIR; s37881; s37881.
DR PIR; s39129; s39179.
DR PIR; s44524; s44524.
DR PROSITE; PS00995; TCP1_3.
KW HYPOTHETICAL PROTEIN.
SQ SEQUENCE 441 AA; 49647 MW; 1018731 CN;

Query Match          4.6%; Score 115; DB 8; Length 441;
Best Local Similarity 27.2%; Pred. No. 4,76e-02;
Matches 25; Conservative 25; Mismatches 37; Indels 5; Gaps 4;

Db      349 ewetllkqeehgsakdnpqektkklmdpcgaqlmmntelpcvsmngptpyppyl 408
       :::|::| |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Qy      242 QIPCYAIRHEHLAGTR-GLEPGSRRITGLADISSGTSYSLPNIQCPSPPTHP- 299

Db      409 pfgiipfpmfpmpfmatlc-nphgadsp 439
       | | :||:| ||: : | |::|:
Qy      300 PTG--QYTLPPLPPTLPPTVQLHLLDPSSA 329

RESULT 15
ID POLR KYMW STANDARD; PRT; 1874 AA.
AC P36304;
DT 01-JUN-1994 (REL. 29, CREATED)
DT 01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)
DT 01-JUN-1994 (REL. 29, LAST ANNOTATION UPDATE)
DR RNA REPLICASE POLYPROTEIN (EC 2.7.7.48).
OS KENNEDYA YELLOW MOSAIC VIRUS (STRAIN JERVIS BAY) (KYMW).
GN VIRIDAE; SS-RNA NONENVELOPED VIRUSES; TYMOVIRIDAE.
RN [1]
RP SEQUENCE FROM N.A.
RC DING S., KEESE P., GIBBS A.;
RA
```

Query Match	4.6%;	Score 115;	DB 6;	Length 1874;
Best Local Similarity	40.9%;	Pred. No. 4.76e-02;		
Matches 27;	Conservative	8;	Mismatches 25;	Indels 6;
			Gaps 6;	

[illegible]

Search completed: Mon Apr 1 11:00:49 1996  
Job time : 25 secs.